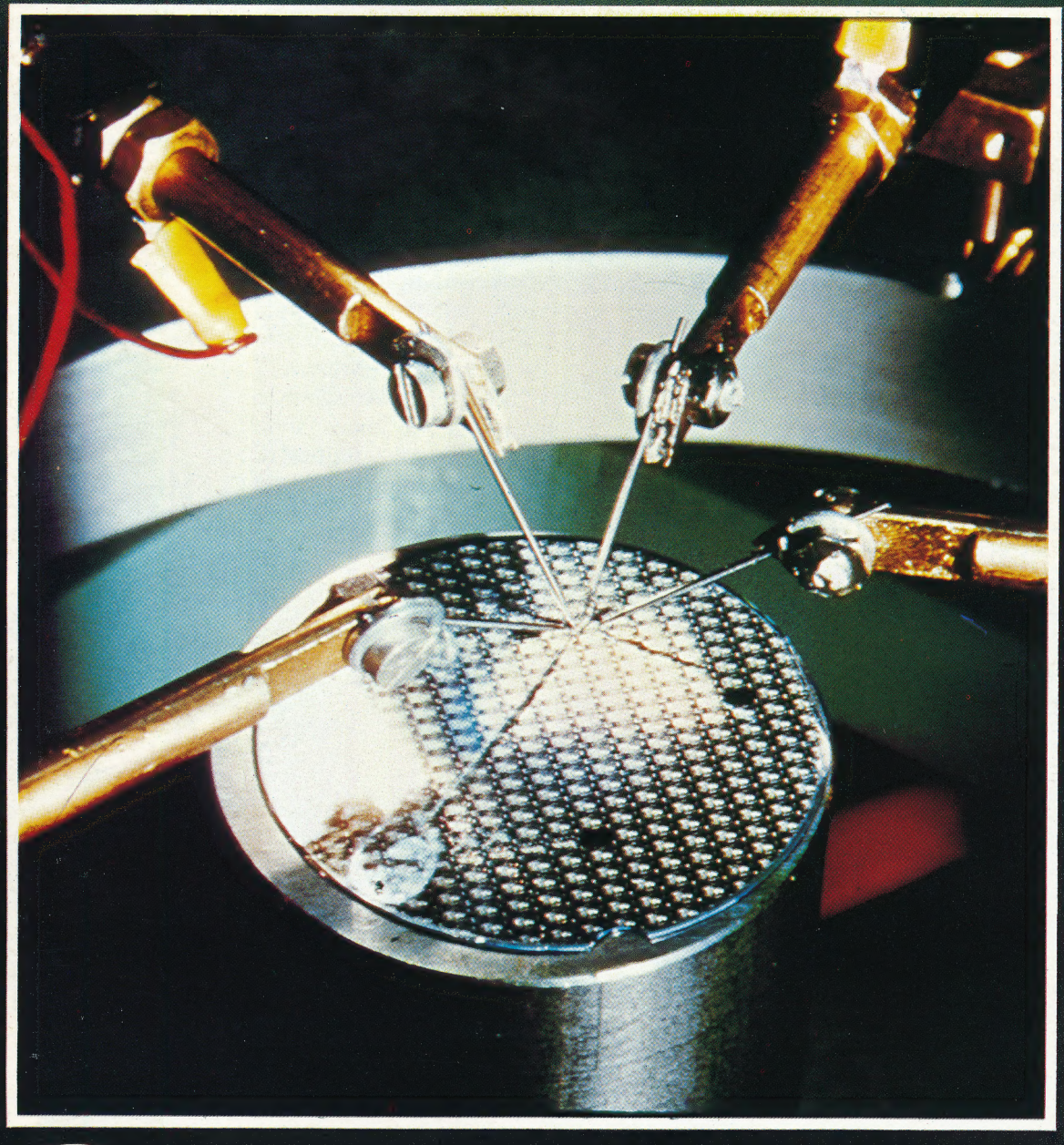
I.T.E.C.



Information
Technology·
Electronics·
Computers·

11 in weekly parts

THE HOME STUDY COURSE IN ELECTRONICS AND COMPUTERS

U.K. 99p Australia $2.50 New Zealand $3.00 South Africa R2.75

Contents Part 11

I.T.E.C. is complete in 51 weekly parts (50 parts + index).

POSTAL SUBSCRIPTIONS

UK readers
You can arrange to receive weekly copies of *I.T.E.C.* at 99p each inclusive of postage and packing. Simply calculate the cost of your subscriptions by multiplying the number of parts you require by 99p and send a cheque or postal order for that amount to:–

Subscriptions Dept., GEJ Publishing Ltd
187 Oxford Street, LONDON W1R 1AJ.

with a note of your name and address, where you wish the copies to be sent and the part number from which you want the subscription to start. Please allow 28 days for your subscription to be processed plus postal time.

Eire readers
As above, but you should calculate the cost of your subscription at IR£1.75 per part.

Other overseas readers
May pay in the currency of their choice provided that its value when converted to sterling is £1.50 per part. Minimum subscription 10 parts.

CONSULTANTS

This series has been produced in collaboration with **TEXAS INSTRUMENTS** Learning Centre. Editorial and graphics assistance is gratefully acknowledged.

U.K. Technical Consultant
Dr. Robert King, Reader in Communications Engineering, Imperial College of Science and Technology, London.

BINDERS

U.K. & Eire
Details of how to obtain your binders for *I.T.E.C.* appear on the back cover.

Australia
Write to *I.T.E.C.* Binders, Gordon & Gotch Ltd., PO Box 213, Alexandria, New South Wales 2015.

New Zealand
Write to *I.T.E.C.* Binders, Gordon & Gotch (NZ) Ltd., PO Box 1595, Wellington

South Africa
Binders are available at R7.95 (please add sales tax) from any branch of Central News Agency. In case of difficulty please write to Intermag, P.O. Box 57934, Springfield 2137.

Other countries
Binders are available from your local newsagent.

BACK NUMBERS

U.K.
Back numbers can be ordered from your newsagent or from *I.T.E.C.* BACK NOS., Whinfrey Strachan Ltd., 187 Oxford Street, London W1R 1AJ. Price 99p each inc. p + p. Please allow 21 days for delivery.

Eire
As for U.K. but please remit in Sterling if ordered by post.

Other countries
Copies available from your local newsagent.

INDEX

The final part of *I.T.E.C.* will be a fully cross-referenced INDEX to the complete work

Editor:
Amanda Kirk

Technical Editor:
Keith Brindley

Art Editor:
Alistair Carlisle

Sub Editor:
Nicholas Bellenberg

Copy Consultant:
Chris Wallace (Creative Technical Services)

Published by GEJ Publishing Ltd.

Distributed by Whinfrey Strachan Ltd.
187 Oxford Street, London W1R 1AJ

Printed in England by Southernprint Ltd.

Photographs, unless otherwise stated, are from the library of Texas Instruments.

Cover: computer controlled chip manufacture. (Photo: IBM).

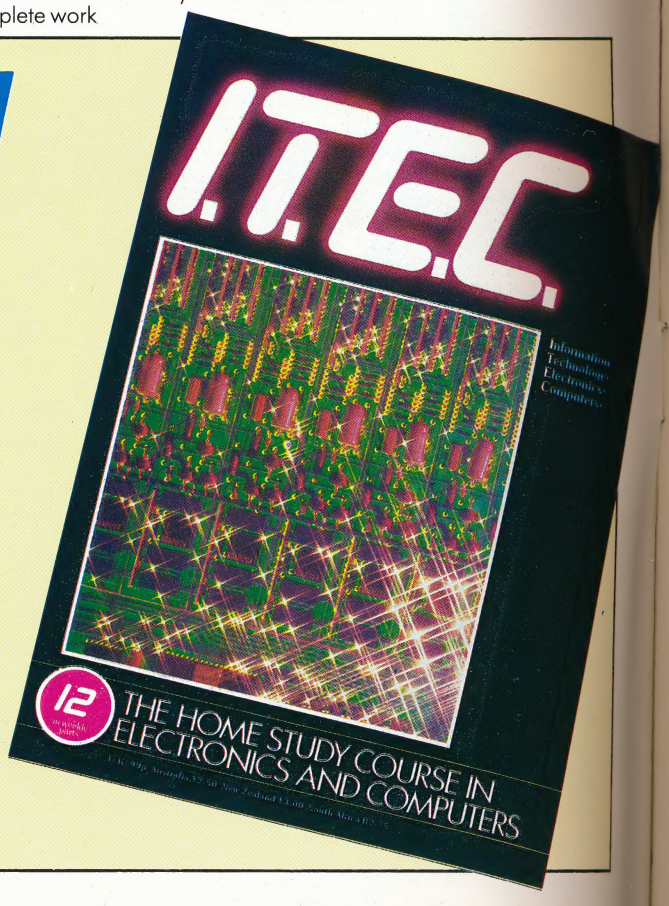
COMING IN PART 12

Continuing *Solid State 11* we find out how to use the **characteristic curves** of a device to determine its usefulness for a specific application.

Computer Science 7 takes a look at the different ways in which **data structures are stored** in mass memory media, such as magnetic tape and disk.

Have you ever wondered why the **digital approach** to building systems is **preferred** over the **analogue approach** for some applications, and vice versa? Find out in *Digital Electronics 9*.

PLUS: *Basic Theory Refresher* – parallel and series reluctances in a magnetic circuit.



10

SOLID STATE ELECTRONICS

Semiconductor manufacture

Building components

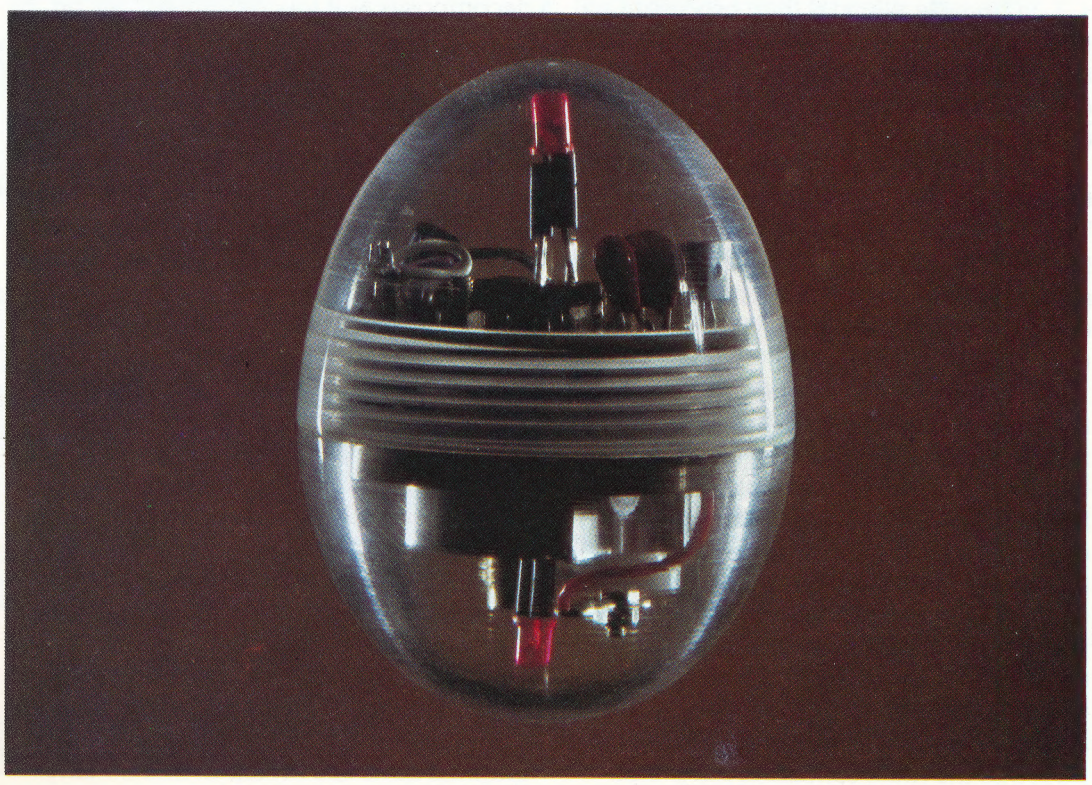
In this chapter we shall be looking at the way transistors and diodes are manufactured. The basic principles also apply to the manufacture of ICs and detailed information about the construction of specific semiconductor devices will be covered in later chapters.

Semiconductor component manufacture can be broken into three basic steps:
1) **preparation** – production of the semiconductor materials which are to form the actual chip at the heart of the component;
2) **mechanical assembly** – the formation of connections from the chip to the component's leads, and packaging;
3) **testing** – component quality control and type determination.
These three steps will now be discussed in greater detail.

Preparation

Most modern semiconductor components begin life as sand (which is composed mainly of silicon). The sand is chemically treated and combined with other materials to form a liquid compound, e.g. **trichlorosilane**, from which a refined **polycrystalline** form of silicon metal is then extracted. (Polycrystalline means many different crystals – even a tiny piece of this material will consist of thousands of silicon crystals.) A similar process was used for the production of germanium crystals for germanium semiconductor components which are not now widely used. Only a few semiconductor component manufacturers have the chemical plant required to produce polycrystalline silicon, other manufacturers buy in their requirements from them.

The next stage in the process is to



Ace Photo Agency/Chris Arthur

Right: electronic egg used to test sensitivity of mechanical handling during component manufacture

produce **monocrystalline** silicon, by a refinement known as **crystal growing**. Monocrystalline silicon comprises, as the name suggests, a single crystal of silicon – whatever size the block of silicon metal.

A quantity of polycrystalline silicon is melted in a crucible about the size of a teacup (*figure 1a*). A small seed of monocrystalline silicon is then lowered into the crucible to just touch the surface of the melted silicon (*figure 1b*). Since the seed is cooler, the molten silicon will crystallize on the surface of the seed, reproducing its monocrystalline structure. The seed is slowly and continuously lifted away from the melt, and more and more of the molten material accumulates to build up a large crystal. The crystal and the pot are rotated in opposite directions as the crystal is pulled upwards in order to retain uniform growth.

The finished product (*figure 1c*), a cylinder of 2-5 inches diameter, contains enough pure monocrystalline silicon to make millions of ICs, transistors or diodes when it is sawn up into tiny chips. All the atoms in each of these chips have to be lined up in exactly the right pattern – forming what is called a **crystal lattice** – to produce a semiconductor.

This pure silicon is then processed so that the p and n-type material is in just the right place to make working semiconductor devices. There are three basic ways of doing this, covered next.

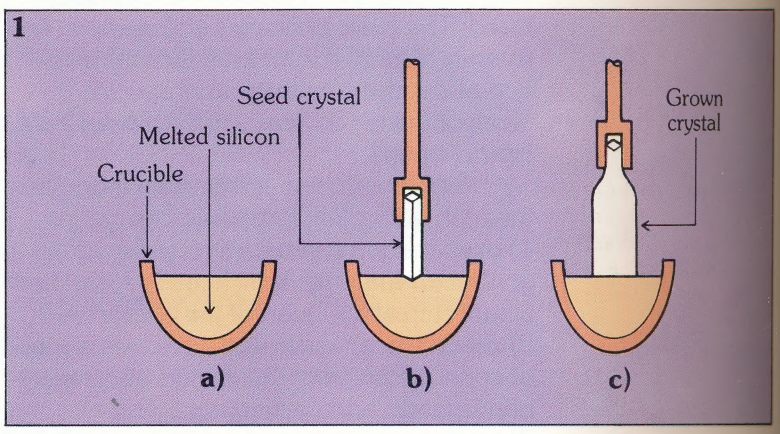
Growing junctions

The first method ever used to make semiconductor components utilised molten silicon. Although this process is not now used, it is worth taking a quick look at it (shown in *figure 1*).

Assume that the molten material in the crucible has been doped to produce n-type material. After some n-type crystal has grown, a pin-head size pellet of p-type impurity is dropped into the molten silicon cancelling out the effects of the n-type impurity and making the crystal grow a region of p-type material. After a precisely timed interval (in the order of a few seconds) – during which time the p-region has grown to about one half thousandth of an inch thick – a pellet of n-type material is added to the melt, cancelling the p-type growth and stimulating renewed growth of n-type material. The overall result is an n-p-n structure.

Currently, the most popular method of growing a crystal is onto a surface from a vapour. This is known as **epitaxial growth**. The first stage involves the growth of a complete silicon bar – doped n-type for example – which is then sawn into wafer thin disks, each ground flat and polished to mirror smoothness.

These n-type wafers or slices are placed onto a heated platform in an enclosed chamber (*figure 2*). Silicon chloride gas is pumped into the chamber and decomposes as it touches the hot silicon slice, leaving a layer of monocrystalline silicon. In this case, it is the gas which contains the p-type impurity, such as boron. The new silicon layer is p-type, making a p-n junction. For transistor manufacture the other region would be added at a later stage by the **diffusion technique**, which we shall look at later.

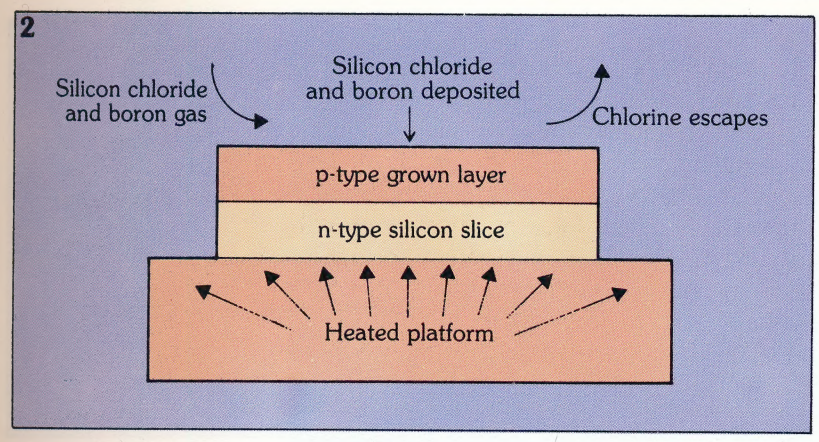


1. The three stages of growing pure monocrystalline silicon.

Junctions from the alloy method

The third basic technique for producing semiconductor junctions is the **alloy method**. Although this has been used to produce billions of germanium transistors and diodes, relatively few silicon semiconductor devices are made this way because of the problems that arise from the high melting point of silicon. However, to avoid confusion we shall look at the manufacture of a silicon alloy diode as an example.

The first step is to grow an n-type crystal which is then cut into slices. The slices are then placed on a flat carrier with graphite disks laid on top of them. Many



2. Epitaxial growth: a p-type layer is grown onto an n-type silicon slice.

3. The alloy method – small holes are bored into graphite disks and tiny balls of aluminium are dropped into them.

4. A p-n junction capped with aluminium.

small holes are bored into these disks and tiny balls of pure aluminium are dropped in them (*figure 3*). Aluminium is a p-type impurity – so in this case the aluminium balls serve as the p-type deposit.

The carriers, with the slices, graphite disks and aluminium balls are placed in a furnace, at a temperature just hot enough to melt aluminium. Some of the melted aluminium dissolves a little of the silicon – the aluminium is said to **alloy** with the silicon.

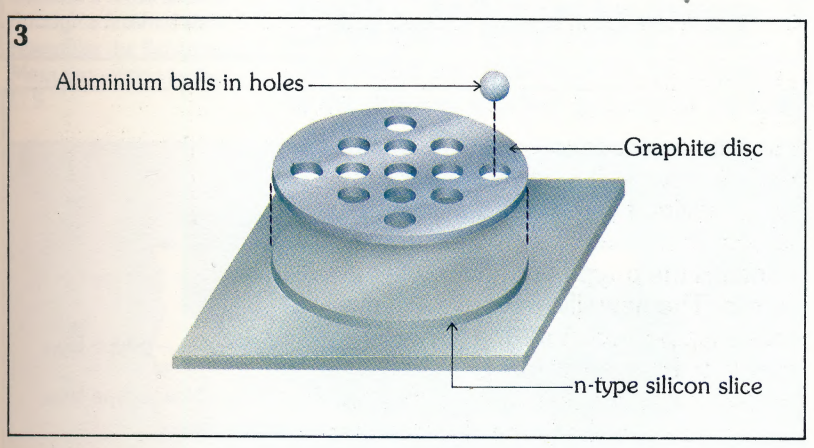
When the whole assembly is taken out of the furnace the aluminium hardens. The graphite disk is removed, leaving pure aluminium on top, aluminium silicon alloy in the middle (the p-region) and the original n-type silicon on the bottom.

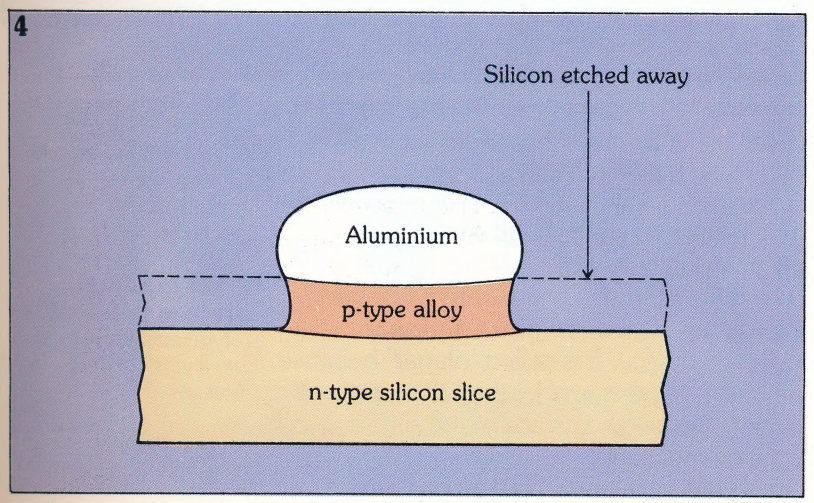
To avoid a short circuit between the contact aluminium and the n-type silicon around the edge, the slice is washed with an acid which dissolves silicon but not aluminium, thus etching away the silicon where it is not protected by the aluminium ball. This is shown in *figure 4*. As you can see, this provides a slice containing a number of p-n junctions, each capped with aluminium.

The individual diode elements are separated and then packaged to produce the finished device. *Figure 5* shows this package – a glass sleeve with cathode and anode leads stuck into each end. The n-side of the semiconductor element is attached to the cathode plug while the p-side is connected to the anode by a whisker of wire. This whole package makes up a silicon alloy diode.

One can see that if this process was repeated on the blank side of an unseparated slice, p-n-p elements of the type used in transistors could be made. However, if the furnace temperature is a few degrees too cool, or if the slice is taken from the furnace a few seconds too soon, then the base region will be too wide. Conversely, if the temperature is too high, or the slice heated for too long, the alloy punches through the silicon wafer and fuses onto the alloy the other side – completely obliterating the base region.

The serious problems involved in the control of factors like these, at the high temperatures needed to work with silicon, are the chief reasons why the alloy process has not been widely used to make silicon transistors. Germanium transistors and certain silicon diodes can, however, be easily manufactured by this process.





Diffused junction transistors

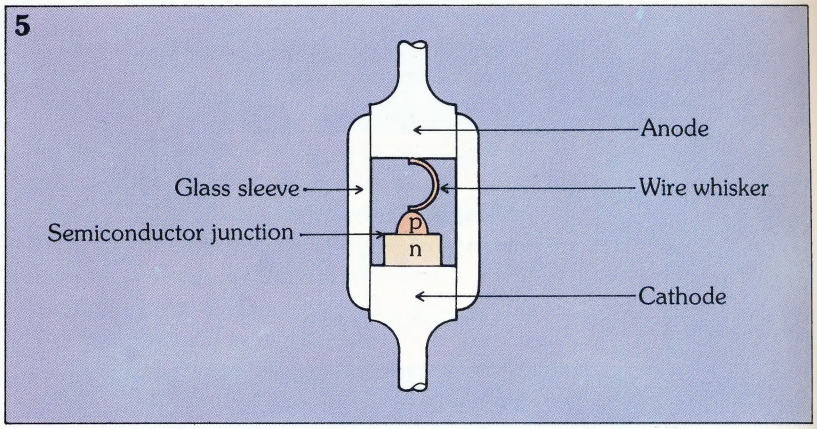
The **diffused junction** process is widely used to make transistors and nearly all integrated circuits. The starting point is again a slice of monocrystalline silicon – in this case it is n-type and is used as the

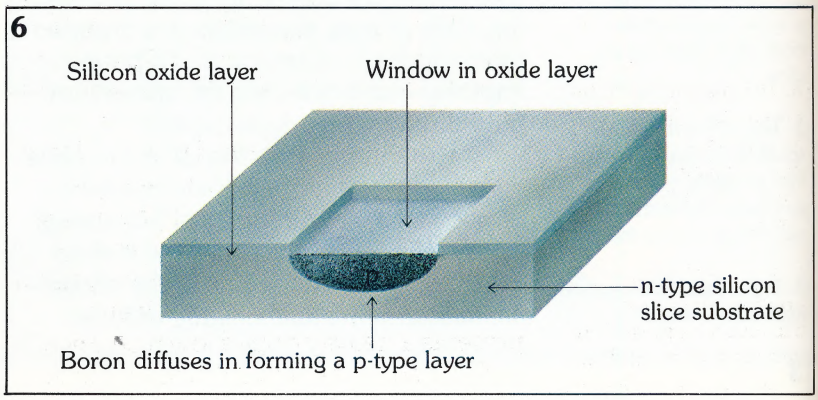
substrate. (Substrate simply means underlayer.) The top of the substrate is oxidized – the oxide layer acting as a protective seal. This layer is, in turn, coated with a chemical called a **photoresist**. The photoresist is sensitive to light – much like ordinary photographic film. A type of photographic negative called a **photomask** is placed on the photoresist and the whole assembly is exposed to light. If a positive photoresist is used, then the light that passes through the parts of the photomask that are transparent causes a chemical change in the photoresist. The slice can now be washed with a suitable solvent, washing away the parts of the photoresist that have been exposed to the light, revealing the oxide layer. The slice is then 'developed' by placing it in a tank of etching chemicals which effectively 'cut' small windows in the oxide, as shown in *figure 6*.

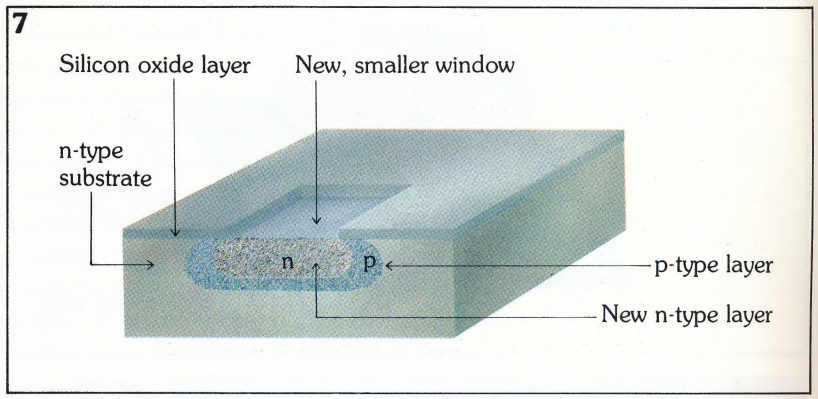
Each slice has hundreds or thousands of windows etched onto its surface. If the slice is to be used for the manufacture of small-signal transistors, each window will be about twenty thousandths of an inch wide. If integrated circuits are being made, the windows may be even smaller than one half thousandth of an inch.

The prepared slice is put into a furnace and heated to a temperature near the melting point of silicon. Boron gas (a p-type impurity) is pumped into the furnace and diffuses into the silicon layer where it is exposed by the windows. Diffusion is a *soaking in* process – like water soaking into a sponge. During a period of several hours, the boron atoms replace a small proportion of silicon atoms in the crystal structure. This has the effect of turning the diffused layers into p-type silicon and creating a p-n junction. When the boron has diffused to the correct depth, the slice is removed from the furnace and the process stops.

A new layer of silicon oxide is then created, covering the top of the slice. New windows are etched so that they lie within the area of the old windows. The slice is then replaced in the furnace and exposed to gaseous phosphorous. The phosphorous atoms diffuse into the window areas and overpower the effects of the boron in a small area within the p-region. This area







reconverts to n-type silicon (see *figure 7*). It's easy to see the transistor structure now – a layer of p-type material between two layers of n-type material.

The planar epitaxial and mesa diffusion processes

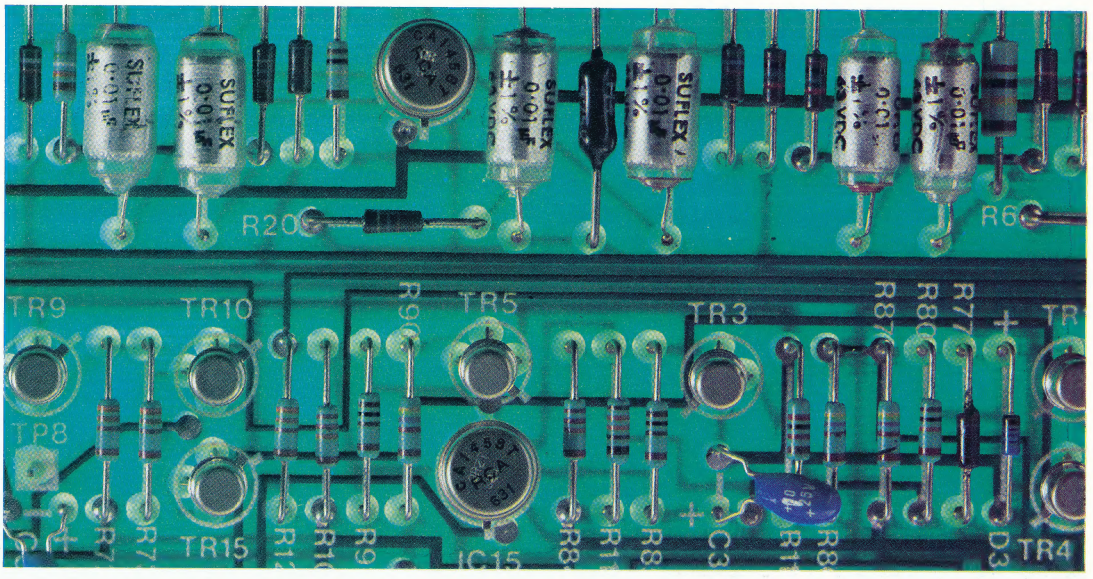
The diffused transistor we have just been discussing is an example of a **planar** diffused device. It is called 'planar' because all of the regions and both of the junctions are on the upper surface of the slice, i.e., all in the same plane.

One variation of the planar process is

5. A finished silicon alloy diode.

6. The diffused junction process – small windows are etched into the oxide layer.

7. The transistor structure – a layer of p-type material between two layers of n-type material.



Tony Stone Associates

Right: a double sided printed circuit board showing a selection of components.

called the **planar epitaxial** process. Here, a lightly doped n-region is grown epitaxially on top of a heavily doped n-type substrate (see *figure 8*). This gives a lightly doped collector region which is very useful as it can handle high voltages. It also presents a heavily doped low resistance path for current to the collector contact. The emitter and base regions are created by the diffused junction process as we have previously explained.

Figure 9 explains how the **mesa** process was so named. In the south west of America flat-topped mountains are called mesas. A cross-section through a finished slice made by this method reveals a line similar to a group of flat topped mountains.

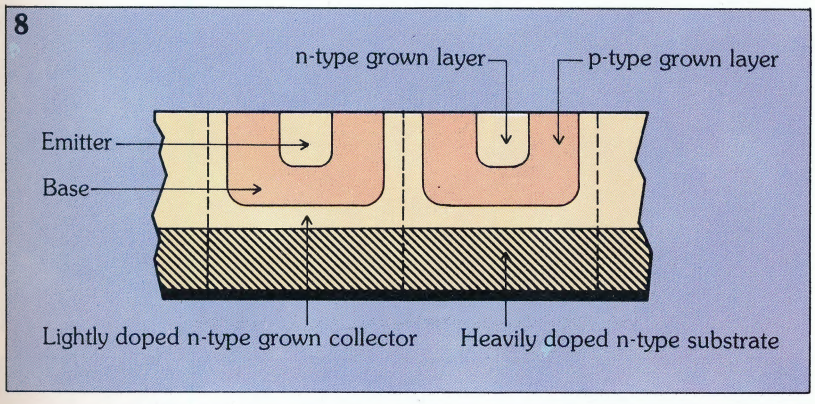
Remember, in the planar diffusion process the p-type base regions were diffused separately through windows in the silicon oxide. In the mesa process on the other hand, a single p-region is diffused or deposited epitaxially over the entire surface of the slice; n-type emitter regions are then diffused individually into the slice – as in the planar process. Finally, to prevent damage to the collector-base junctions when the chips are sawn apart, acid is used to etch depressions between the individual devices which are then sawn apart in the places indicated by the dotted lines in *figure 9*.

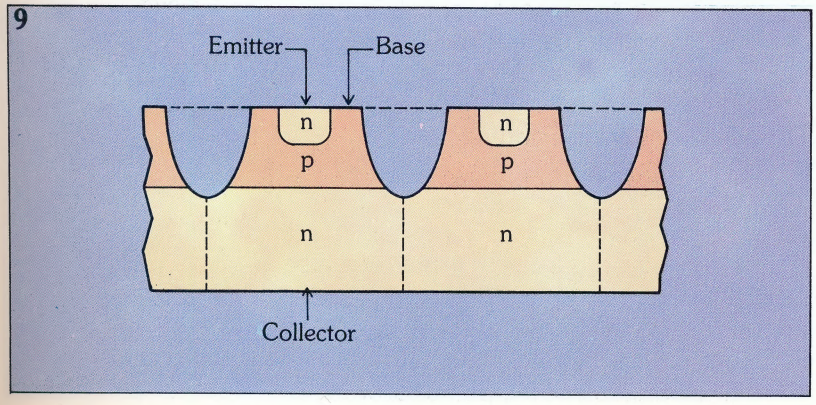
One advantage of the mesa over the planar process, is that it produces fewer defects per slice. However, the planar process is inherently cheaper.

We won't go into all the advantages of the various processes here, suffice to say that a semiconductor designer must assess performance, reliability and cost when deciding which process steps to use. It is a measure of the importance which designers place upon using the right processes to obtain the desired component, that computers are often used to select the best combination.

8. The planar epitaxial process produces transistors with a lightly doped collector region – ideal for handling high voltages.

9. Mesa diffusion process, so named because a cross-section through a finished slice resembles the flat-topped Mesa mountains in the U.S.



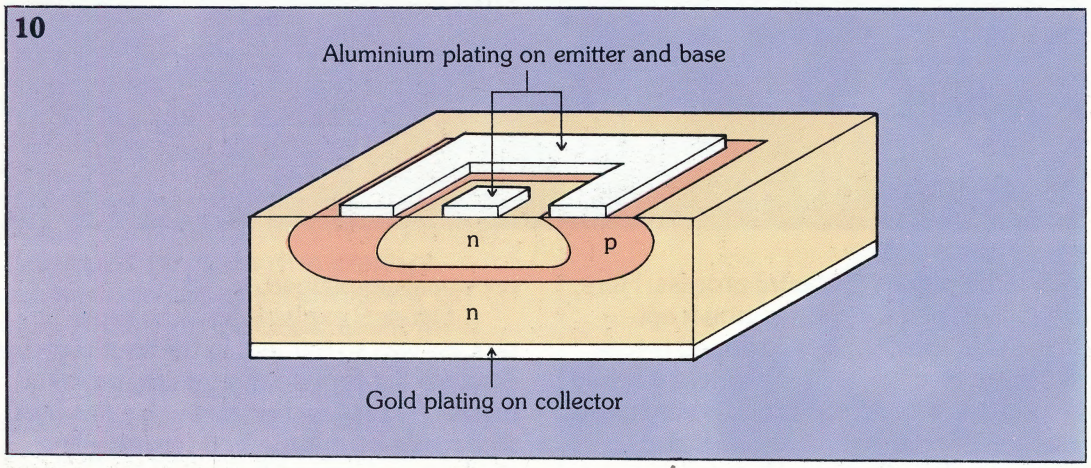


Mechanical assembly

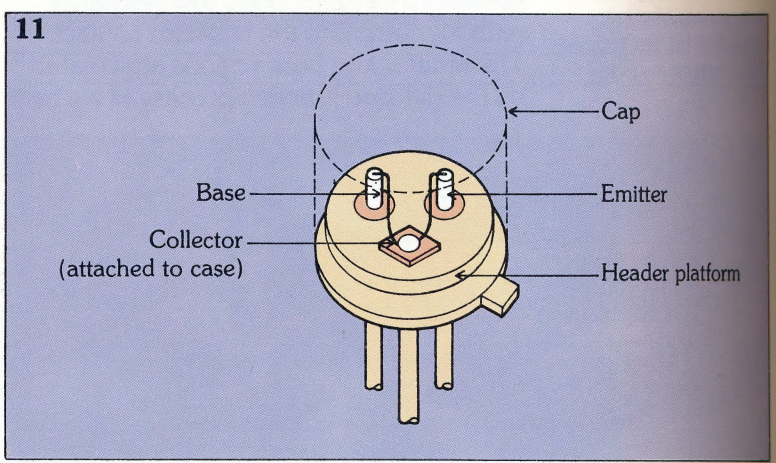
The second stage in building transistors is to connect the three layers of the n-p-n (or p-n-p) junctions to the terminal leads of the transistor. In order to do this the surface of each region has to be plated with a thin layer of metal. The same sort of photomasking techniques that we have recently looked at are used to do this. As you can see in *figure 10*, the base and the emitter regions are plated with aluminium, while the collector – the bottom of the slice – is plated with a layer of gold. The gold plating allows a good electrical connection to be made, as well as providing an excellent solder contact – used when the chip is bonded to its package. Soldering ensures that a strong physical bond is made over a wide area, allowing heat to be dissipated, via the package, to the environment.

The last stage in the assembly of a transistor is its packaging. *Figure 11* shows how a diffused chip is packaged in a metal body or can. The chip is fastened to the gold-plated header platform by a heat process which, in effect, solders the chip to the gold-plating. This platform usually serves as the collector contact and is directly connected to the terminal lead. The emitter and base leads are insulated from the can by glass – shown as the grey rings in the drawing. The base and emitter sections of the chip are connected to the leads by thin gold wires – about the diameter of a human hair. A metal cap is then welded onto the header. The transistor is now complete and ready for testing.

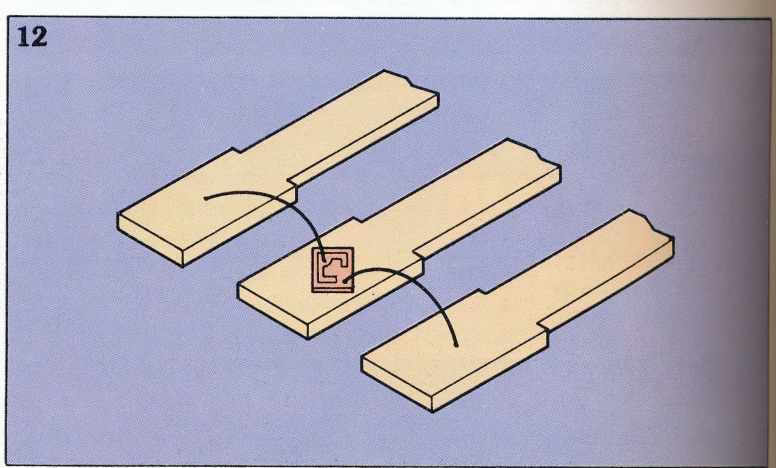
An alternative method of packaging is to encapsulate the transistor in a solid block of plastic. Typically, the three leads – flattened as shown in *figure 12* – are clamped in a temporary fixture. The chip is soldered to the centre lead (collector) and gold wires make the connections to the base and emitter. This assembly is placed in a mould which is filled with liquid plastic. When the plastic has set we are left with a very rugged transistor.



10. Each n or p-region is plated with a thin layer of metal providing electrical connections to the terminal leads.



11. A diffused chip packaged in a metal can.

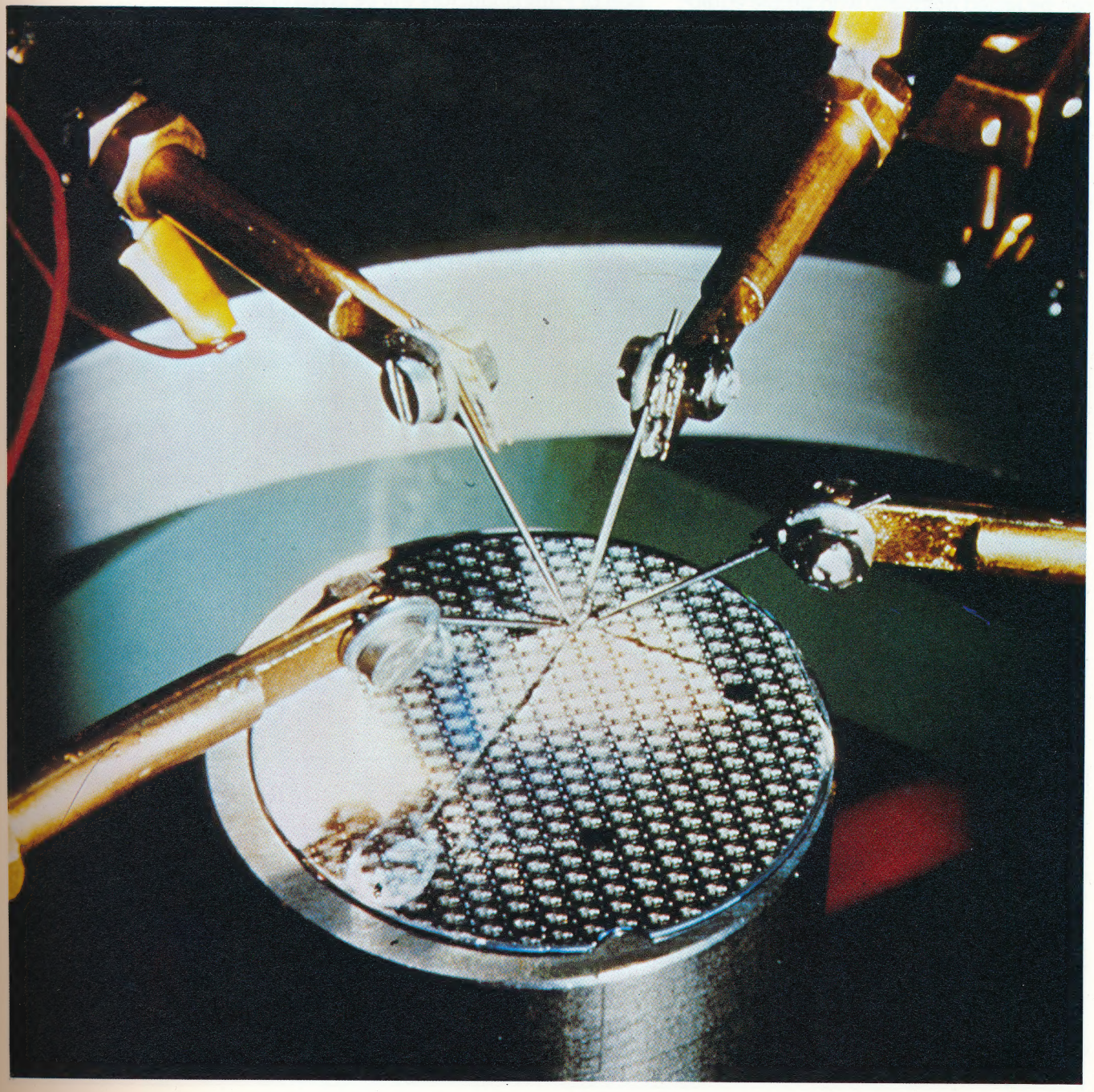


12. An alternative method of packaging the finished transistor – in a block of plastic.

Testing

The third and extremely important step in semiconductor manufacture is testing. In most modern industries, random samples of finished products are tested: it is assumed that untested products from the same batch will function in the same way. In the semiconductor industry, however, every device must be tested and **characterised** before its type number can be determined. Strange as it sounds, even in an industry as advanced as this, it is still necessary to test each device to find out what's been made! Despite using some of the world's most sophisticated techniques and experience, few batches of semiconductors produce components with identical electrical characteristics.

Each component is tested by connecting it to a test circuit. Depending on test results, components from a single batch may be divided into, say, four or five different categories and sold under diffe-



Below: computer controlled chip manufacture. (Photo: IBM).

rent type numbers. Components that do not meet the specification of any of these categories are rejected.

The specification and performance figures of a semiconductor component are given on the manufacturer's data sheet. Testing and characterisation ensure that the component is within specified limits of forward voltage, reverse current and many other characteristics that you will be familiar with from earlier chapters on diodes.

Although semiconductor components can be tested manually with simple equipment, manufacturers generally use sophisticated automatic systems for testing and grading. These can handle thousands of different components every hour, automatically testing and measuring many characteristics, analysing the results to determine component type and depositing the tested components in the appropriate hoppers. Many components are also subjected to operating life tests – in which transistors can be actually operated day and night for weeks, often in high temperature ovens, to ensure that they will perform reliably in adverse conditions.

Further information on the production of semiconductors will be given in later parts. In the next chapter we shall begin to discuss the electrical characteristics of transistors in detail.

Glossary

alloy method	a way of making p-n junctions in which aluminium (a p-type impurity) is alloyed to an n-type semiconductor material
crystal growing	process by which polycrystalline silicon is refined into monocrystalline silicon. A large crystal is grown from molten polycrystalline silicon onto a monocrystalline seed. The silicon that crystallizes onto the seed reproduces the seed's structure
diffused junction process	method used to make p-n junctions in diodes, ICs and transistors. The substrate is selectively etched away and exposed to dopants in the form of gas. The impurities in the gas diffuse into the substrate altering its chemical composition and making a p-n junction
epitaxial growth	process by which a p-type layer of silicon is grown onto an n-type wafer, by exposing it to silicon dichloride gas
mesa diffusion	process that diffuses p and n-type layers together and then etches away channels between the individual components. So called because the transistor chips look like flat topped (mesa) mountains
monocrystalline silicon	pure silicon with a single crystal structure
planar diffusion	method of making transistors that ends up with the base, collector and emitter in the same plane
polycrystalline silicon	pure silicon with many separate crystals in its structure

ELECTRICAL TECHNOLOGY

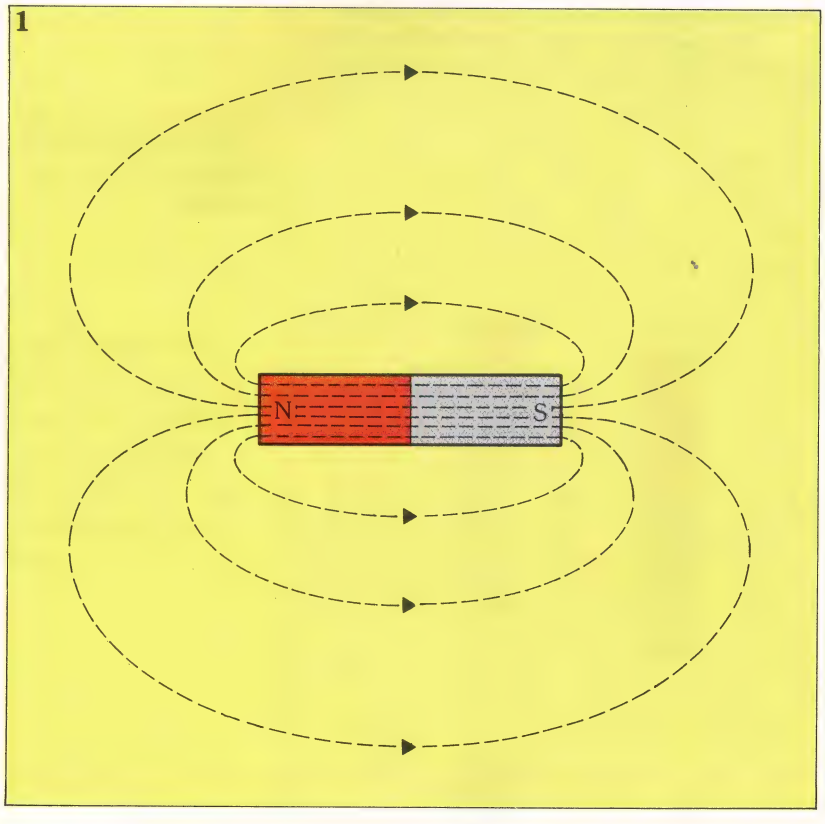
Magnetic flux and permeability

In the last *basic refresher* article is *Part 10* we saw that the flux density of an electric field is proportional to the field strength – the constant of proportionality being the permittivity. Magnetic flux density is related to the magnetic field in a similar way. The ratio of magnetic flux density to magnetic field strength is defined as the **permeability** (symbol μ), which is defined:

$$\mu = \frac{B}{H}$$

where μ is measured in henrys per metre (H/m, which can also be written as Hm^{-1}); the flux density, B, is measured in webers per square metre ($Wb\ m^{-2}$); and the field strength, H, in ampere-turns per metre ($At\ m^{-1}$). (Webers per square metre are also known as teslas, T.)

1. Lines of magnetic flux set up by a permanent magnet.



The permeability of a material tells us how the magnetic effect of a coil is altered by the presence of that material. If a coil is wound in air, the permeability μ has a value of $1.25664 \times 10^{-6}\ Hm^{-1}$. This is actually the permeability of a vacuum, however the permeability of air differs from that of a vacuum by a negligible amount.

The ratio of the permeability of a material to the permeability of air, is termed the relative permeability of the material, μr, defined as:

$$\mu_r = \frac{\mu}{\mu_o}$$

Some materials, for example **ferromagnetics**, have a permeability which is very much greater than that of air – often by up to several thousand times. Most of these materials are composed of iron, nickel and cobalt or some of their alloys. Nowadays, a large number of synthetic ceramic ferromagnetic materials exist, which are sometimes termed **ferrimagnetics**. The permeability of most ferromagnetic materials is not constant, but varies with the field strength – becoming smaller as the flux density increases.

Most other materials have a relative permeability which is constant and independent of the flux density. The value of the relative permeability for these materials is very close to one (or unity). Those materials which have a relative permeability that is slightly greater than unity are termed **paramagnetic** while those with a value slightly less than one are termed **diamagnetic**.

Lines of magnetic flux

We have seen that lines of magnetic flux are continuous loops that are linked to the system that produces them. When lines of flux are set up by a permanent magnet, they appear to start from the magnet's north pole and end at its south pole. However, even here the flux lines form continuous loops, closing back on themselves, passing along the length of the magnet (*figure 1*).

We can now state some properties of lines of magnetic flux:

1) The direction of a line of magnetic flux is that in which the north pole of a compass needle will point when placed on the line.
2) Each line of magnetic flux forms a closed loop (*figure 1*).
3) Lines of magnetic flux never cross. This must be so since, if we place a compass at the point where two lines of flux cut, it would have to point in two directions at the same time.
4) Lines of magnetic flux act like elastic bands – they try to contract in length and so get fatter.
5) Magnetic flux lines lying side by side and in the same direction tend to repel and spread further apart.
6) Where lines of magnetic flux are close together, the flux density is high; where they are further apart the flux density is low.

These properties of flux help us assess the behaviour of magnetic fields.

Mechanical effects of magnetic fields

If we take two magnetic systems and slowly bring them together, we should expect that the lines of flux would either squash together or join up, depending on whether or not the magnetic systems are repelling or attracting.

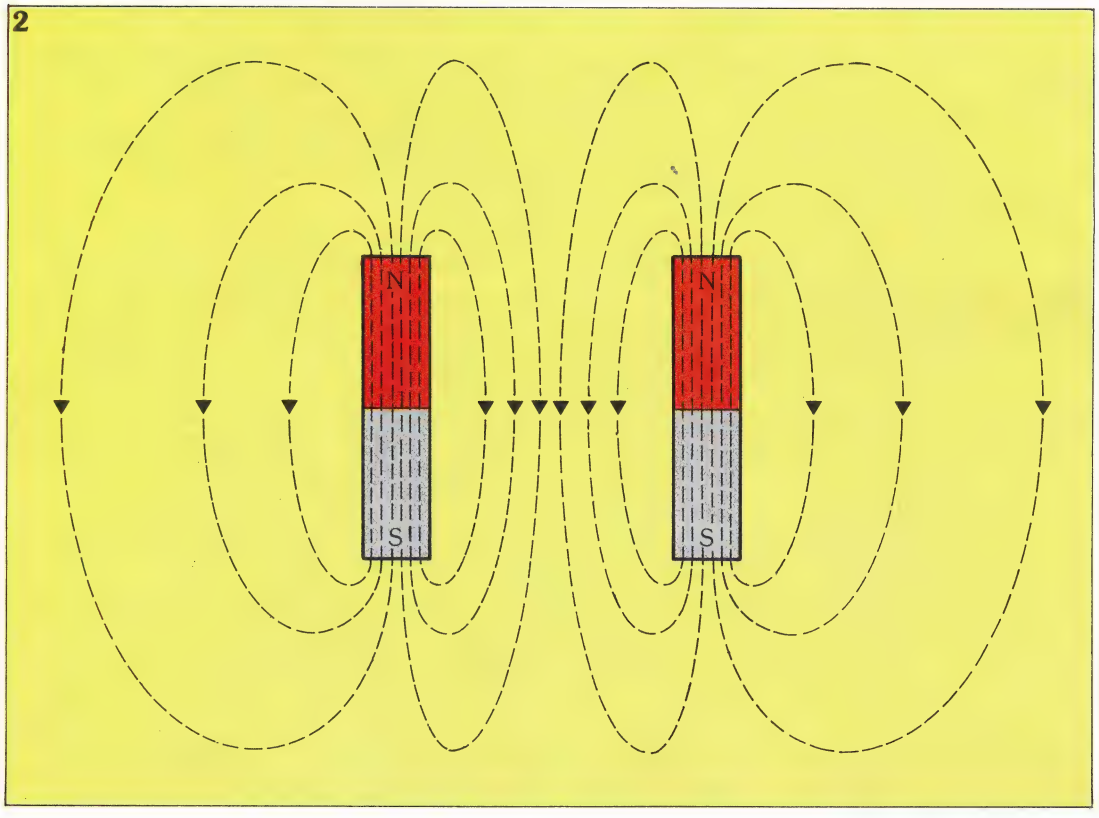
Figure 2 shows two bar magnets parallel to each other with their like poles opposite. Since the lines of magnetic flux are tightly packed between the two magnets they will repel each other. As they move away from each other (because of the force of repulsion) the lines of magnetic flux will become more widely spaced and so the force between the two magnets will diminish. The force between two magnets is inversely proportional to the square of the distance between them.

If the magnets are arranged so that their unlike poles are opposite, the individual field patterns will coalesce when they are close together (*figure 3*). Some of the lines of flux starting from the north pole of one magnet, will cross to the south pole of the other magnet, pass along its length to the north pole and then return to the south pole of the first magnet – completing its loop at the north pole where it started. These lines of flux will tend to contract in length, resulting in a force of attraction between the magnets.

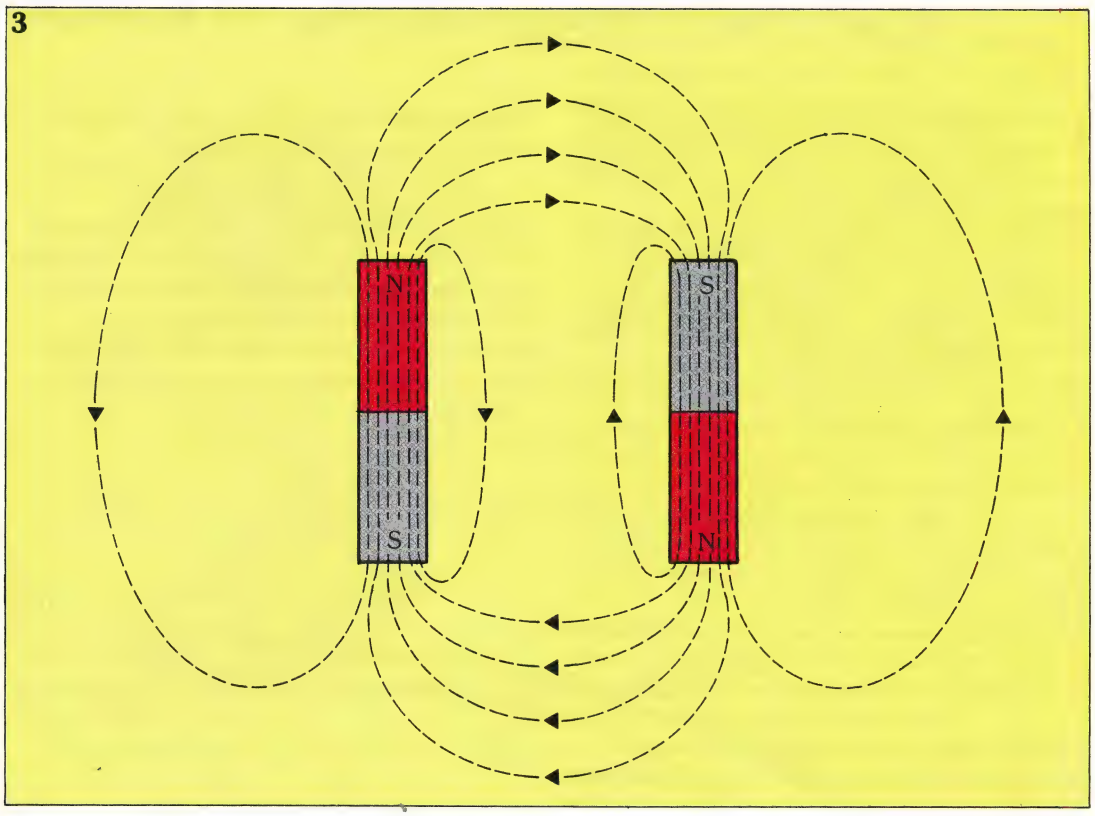
We can see the same interaction between magnetic fields when one of the sets of lines of flux is produced by a magnet and the other by a current flowing in a wire. In *figure 4a*, two poles of a magnet are producing a downward-heading magnetic flux. At some distance to this is a wire, seen to be perpendicular to the page, carrying a current flowing away from us. This creates a magnetic flux in clockwise circles around the wire.

If the wire is placed between the poles of the magnet then these two flux patterns will interact to produce that shown in *figure 4b*. The lines of flux to the right of the wire have increased in number and are packed more closely together. This is because the lines of flux going down from the wire have combined with those from the magnet. There are less lines of flux to the left of the wire, because the lines going up from the wire have cancelled out some of those going down from the magnet.

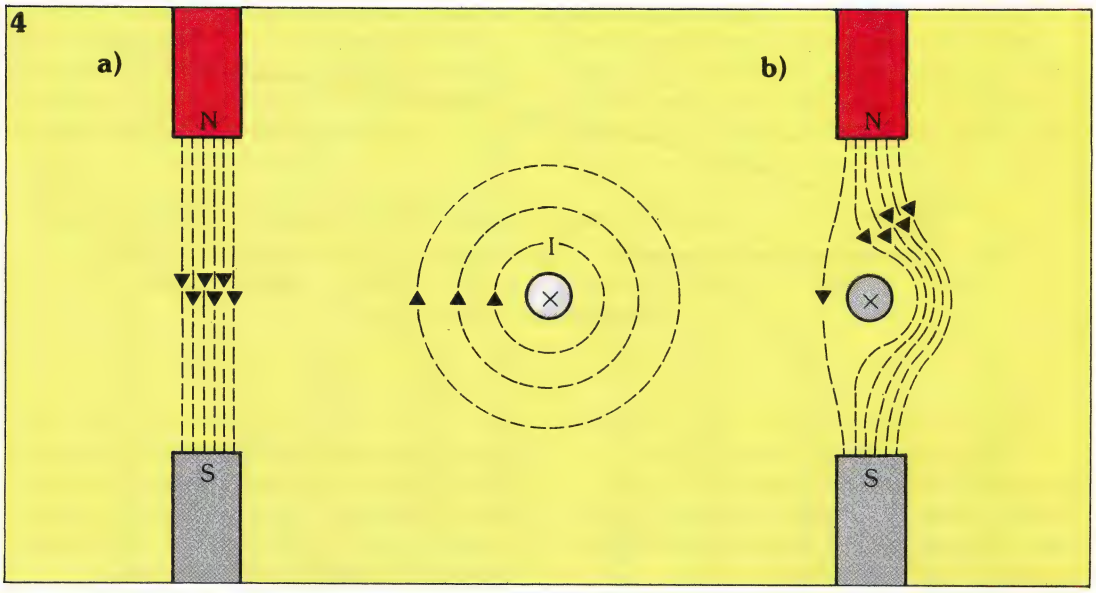
As you might expect, there will be a force exerted by the magnet which will tend to push the wire to the left – away from the region where the lines of flux are dense. The magnet will similarly experience a force – to the right. If the magnet is fixed, however, and the wire is loosely suspended and therefore free to change position, the result will be movement of the wire in a leftward direction. This phenomenon, of course, forms the basic principle behind most types of electric motors.



2. Lines of magnetic flux around two parallel bar magnets with like poles opposite.



3. Lines of magnetic flux around two parallel bar magnets with unlike poles opposite. Note how the flux lines coalesce.



4. (a) Magnetic flux lines between two poles of a magnet and around a wire carrying current, placed some distance apart; **(b)** the flux line pattern when the wire is placed between the two poles.

Energy in a magnetic field

Magnetic fields can store energy, and the energy stored in one cubic metre of a magnetic field, W, can be written in terms of the magnetic field strength, H, and the magnetic flux density, B:

$$W = \tfrac{1}{2} B \times H$$

This energy is measured in Jm^{-3}

If, for example, a rectangular box is placed parallel to the magnetic field then the above formula, multiplied by the volume of the box, gives the stored energy in the box. If we imagine that the box has a length of l metres and a cross-sectional area of A square metres then its volume is given by $A \times l$. The energy stored in this volume is, therefore:

$$W = \tfrac{1}{2} B \times A \times H \times l$$
$$= \tfrac{1}{2} M \times \Phi$$

where M is the magnetomotive force in ampere turns (At) and Φ is the flux in webers (Wb). □

ELECTRICAL TECHNOLOGY

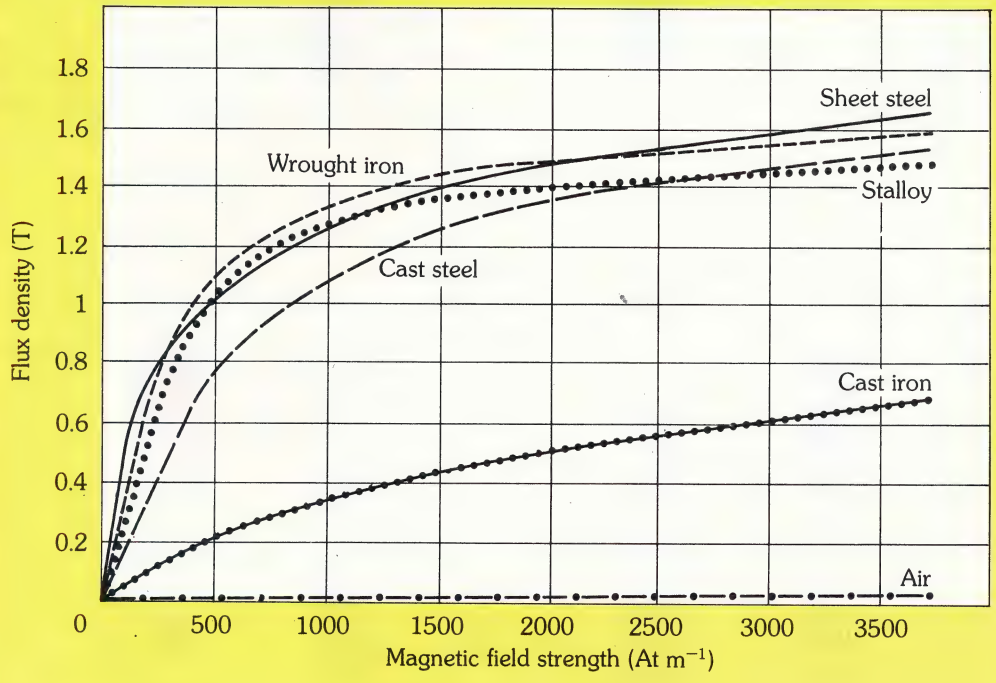
Ferromagnetic materials

Ferromagnetic materials are fundamental to magnetic systems; it is therefore important to analyse their performance – comparing their properties to those of non-magnetic (paramagnetic) materials. We have already seen that the permeability of paramagnetic materials remains constant for all values of the applied flux density. However, this is not true for **ferromagnetic** materials.

fairly constant value. At this point, the steel is said to be **magnetically saturated.**

This behaviour can be explained by a consideration of the way in which ferromagnetic material becomes magnetized. Such material may be thought of as comprising many tiny regions, called **domains:** these are the most basic magnetic elements of a substance. The elementary molecules of the

1



1. The relationship between magnetic field strength and flux density for a number of different materials.

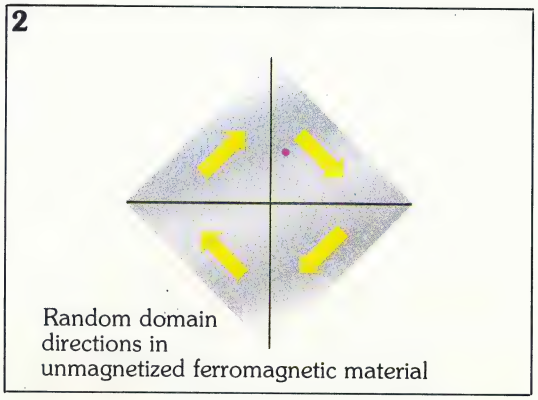
The graph in *figure 1* illustrates the relationship between magnetic field strength and flux density for different materials. This is known as the **magnetization characteristic** of a material. We can see that for a given magnetic field strength, say 3000 At m^{-1}, the magnetic flux density of air is about 0.004 T, while that of sheet steel is 1.6 T – 400 times greater. This is to be expected as steel is magnetic and air is paramagnetic.

We can also see that the magnetic characteristic of air is represented by a straight line, i.e. that the magnetic flux density is directly proportional to the field strength. The flux density for sheet steel, on the other hand, rises quickly at first and then flattens out to a

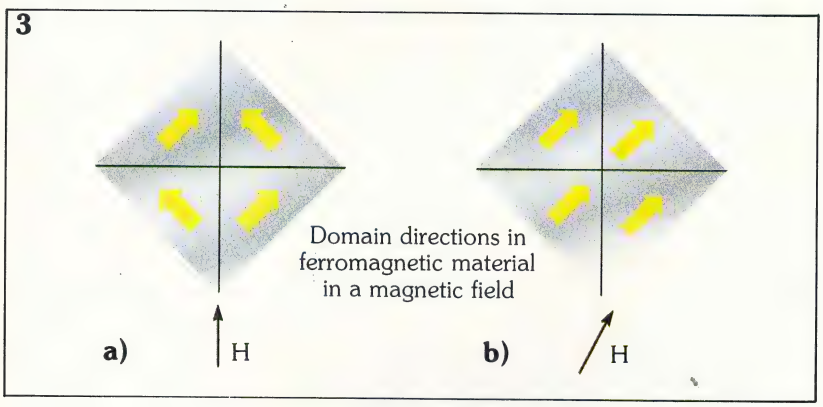
ferromagnetic material, created by the spinning action of electrons around the atom's nucleus, form the domains. The orbital electrons within an atom not only revolve around the nucleus but also spin round on their own axis. In paramagnetic materials, the number of electrons spinning in a clockwise direction is equal to the number of electrons spinning in an anticlockwise direction – the opposite electron spins cancel each other out. Magnetic materials, on the other hand, have more electrons spinning in one direction than the other, resulting in small magnetic effects known as **magnetic moments.**

When ferromagnetic material is in an unmagnetized state, the domains are randomly

2. Individually magnetized domains which cancel each other out.



3. When a magnetic field is applied, (indicated by the arrow) the domains change their direction of magnetism to lie nearly parallel to the applied magnetic field.



oriented, so the magnetization of the domains cancels each other out. The overall effect is that, although the individual domains remain magnetized, the material as a whole appears unmagnetized. This is shown in a very simplified form in *figure 2*. The arrows represent four magnetized domains within an unmagnetized sample of material. Although in reality the domains are randomly oriented, they have been given preferred directions, parallel to the boundaries of the specimen.

Figure 3 shows the application of a magnetic field to this specimen: the direction of the magnetic field is shown by the arrow. As you can see in *figures 3a and 3b*, the domains change their directions of magnetism, so that they lie along the preferred direction most nearly parallel to the applied magnetic field. This occurs as the magnetic field strength is increased from zero. It should be noted that although not all domains switch direction at the same time, most of them flip round soon after the application of the magnetic field. This accounts for the sudden rise in the flux density in the first part of the magnetization curve shown in *figure 1*. After this initial rise, most of the domains are aligned as near as possible along the direction of the magnetic field. Any further increase in the magnetic field has little effect on increasing the flux density.

Values of field strength and permeability for a number of common materials are given in *table 1*.

Hysteresis loops

The magnetization curve in *figure 1*, shows how the flux density varies with the magnetic field strength if we begin with an unmagnetized piece of iron. In most practical applications, the

Table 1
Values of field strength (H) and permeability (μ_r)

Flux density	Wrought Iron		Cast Iron		Sheet Steel		Stalloy*		Air
B (T)	H	μ_r	H	μ_r	H	μ_r	H	μ_r	H×10³
0.1	70	1,140	200	400	45	1,775	80	1,000	80
0.2	90	1,770	450	355	50	3,190	100	1,600	160
0.3	100	2,390	800	300	60	4,000	125	1,920	240
0.4	120	2,650	1,300	246	70	4,550	145	2,200	320
0.5	140	2,850	2,000	200	90	4,420	160	2,500	400
0.6	170	2,810	2,800	171	130	3,670	180	2,650	480
0.7	220	2,540	4,000	140	170	3,290	200	2,800	560
0.8	270	2,360	5,500	117	230	2,770	250	2,550	640
0.9	320	2,240	8,000	90	330	2,190	310	2,320	720
1.0	400	2,000	11,000	73	470	1,700	400	2,000	800
1.1	500	1,750	15,000	58	630	1,390	500	1,750	880
1.2	620	1,545	20,000	48	800	1,195	700	1,370	960
1.3	850	1,230	—	—	1,050	990	1,200	867	1,040
1.4	1,200	930	—	—	1,350	825	2,300	487	1,120
1.5	2,000	600	—	—	1,800	665	4,000	300	1,200
1.6	3,500	370	—	—	3,200	413	7,500	171	1,280
1.7	6,000	336	—	—	5,300	263	14,000	97	1,360
1.8	10,000	144	—	—	9,000	160	24,000	60	1,440
1.9	16,000	95	—	—	14,800	103	—	—	1,520
2.0	25,000	64	—	—	30,000	53	—	—	1,600

**A steel/silicon alloy (3.5% silicon).*

magnetizing force is initially increased from zero to a maximum value, then decreased through zero to the same maximum value in the opposite direction – the whole process being repeated cyclically. The curve that shows the relationship between flux density and field strength, for repeated reversals of magnetic field is called a **hysteresis loop** (*figure 4*).

If you imagine that we apply a magnetic field of magnitude H to a piece of unmagnetized ferromagnetic material, the flux density will be B_{max} – shown by point P on the graph. The initial magnetization curve for this operation is the part of the graph marked with a single arrow.

If we now reduce the magnetic field to zero, the flux density will only fall to point M on the graph. This is because only some of the domains can relax their directions of magnetization. This residual flux is known as **remanence.** If we next increase the field strength in the opposite direction, the flux density will continue to fall until it becomes zero at point A. At this point all the domains have arranged themselves randomly, so that the net magnetization is zero.

The magnetizing force required to demagnetize the material completely is known as the **coercivity** or **coercive force.** The material will only remain completely demagnetized as long as the magnetizing field of value OA exists. If this is removed, the flux will rise again – almost to its earlier value given by OM. The only way we can completely demagnetize a piece of material is to return its properties to the point O.

If the magnetizing force is increased to its maximum negtive value, the flux density will reach its maximum negative value – which corresponds to point P on the graph. Changing direction of the field strength from $-H_{max}$ to H_{max} will take the graph along the path P′M′A′P, completing one full cycle of operation.

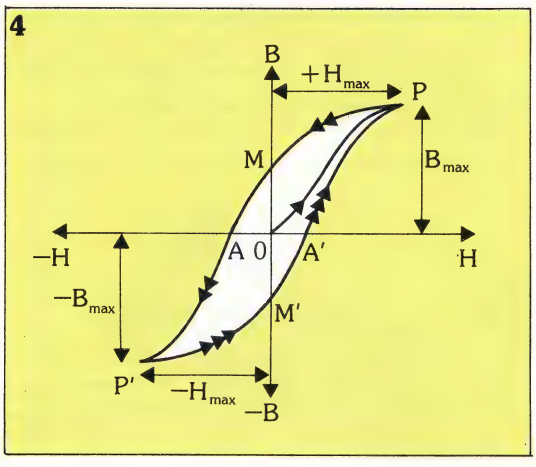
Energy lost in magnetization

Many electrical systems use an alternating current supply which changes its magnitude and direction many times a second, for example, at a frequency of 50 Hz in the National Grid electrical system of the U.K. If we use an alternating current to provide a magnetic field to magnetize a piece of iron, it is obvious that the magnetic field will also change direction at the same rate as the current. The piece of iron will therefore be taken around its hysteresis loop with each cycle of the supply current.

Every time a domain's direction of magnetism is changed, the orientation of the electron orbits is changed. This is done against some internal atom friction and we should therefore expect some energy to be expended when this occurs.

We can calculate the energy dissipated by any material, of volume 1 m^3, moving around its hysteresis loop, by determining the area of the hysteresis loop: the more narrow the hysteresis loop, the less energy that will be wasted. Materials with narrow hysteresis loops, e.g. wrought iron, are used in the construction of transformer cores so that energy loss is minimised. Materials such as these are known as **soft** magnetic materials.

When permanent magnets are manufactured, we need to use materials which have as large a residual magnetism as possible – like steel. These are known as **hard** magnetic materials and have wide hysteresis loops. □



4. Hysteresis loop showing the relationship between flux density and field strength for repeated reversals of magnetic field.

Selection and arithmetical circuits

Data selection units

In the last chapter we looked at the decoder circuit – the second type of structural unit that we shall examine is the **data selection unit.** The function of these combinational logic circuits is to transfer data from various sources to different destinations.

In *figure 1,* a data selector that is similar to the decoder in *figure 30* from the last chapter is shown. As the switch on the right of the diagram suggests, this data selector connects the output line E with one of the four input lines A,B,C or D. In this way, digital information present in any one of these inputs can be routed towards the output line. The input signal is selected by means of the two-input selection control wires W and X. Here the combination – 01 – on the W and X lines, selects input B. The data coming from this gate will be the same as that present at its input, and all the other gate outputs will be logic 0. Thus, the OR gate output, E, will supply the same data that is present at input B.

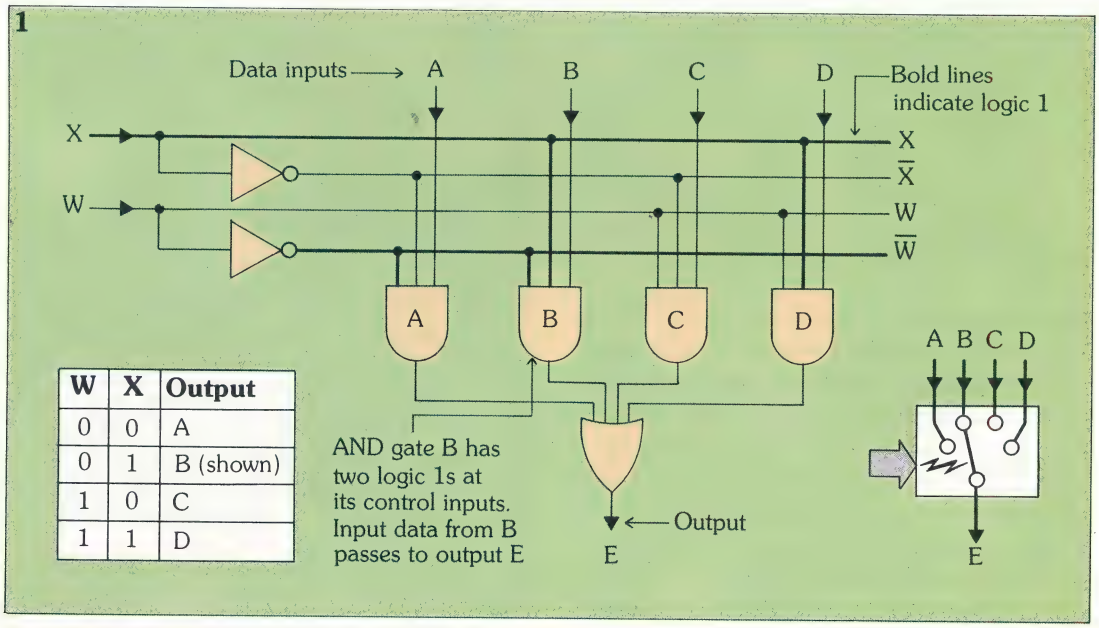
Two control inputs handle the four combinations necessary for selection here; three control inputs would handle eight data inputs; four would handle sixteen data inputs and so on.

How a data selector is used

Two-input data selectors that have only one control line are often used. An example of this is shown in *figure 2:* two 4-bit numbers, A and B, are either routed directly from the display and operand registers to the X and Y inputs of the adder-subtracter, or are made to 'cross over' to the opposite inputs, A to Y and B to X, enabling subtraction of B from A, or A from B.

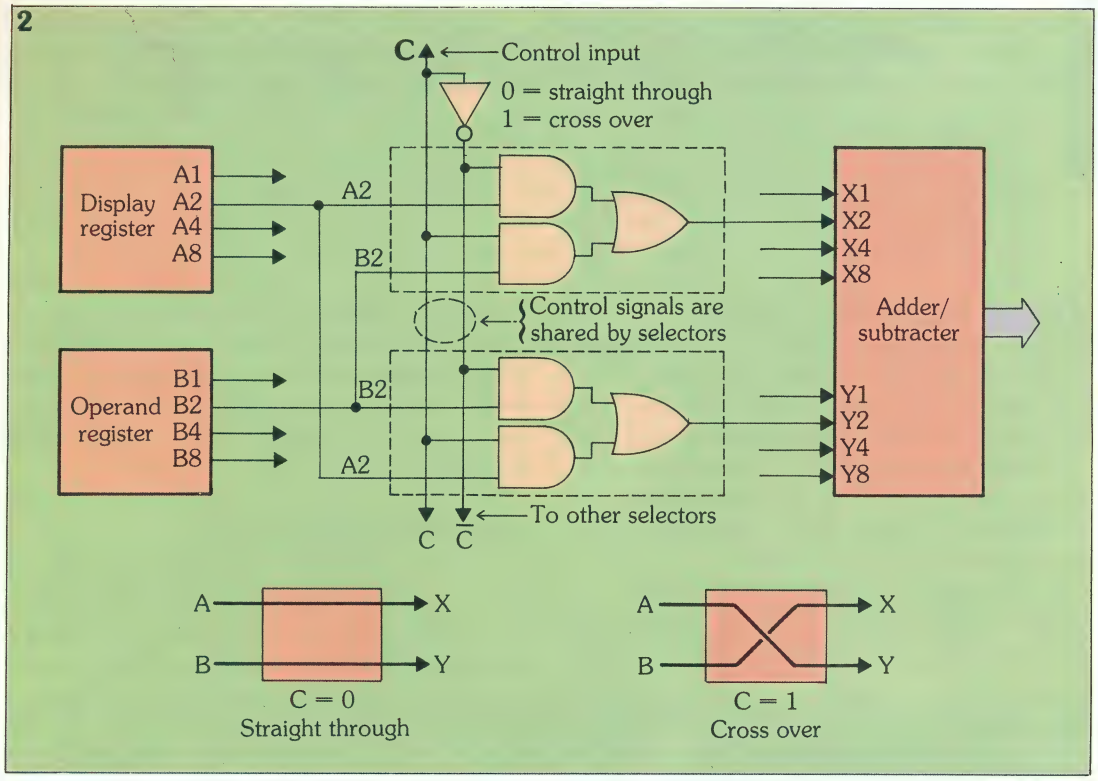
In order to perform these operations, a two-input data selector is needed on each of the eight input wires to the adder-subtracter. (Eight data selectors are needed in total, but only two are shown here.) We need only one control input – labelled C. If a logic 0 is applied to C the

1



W	X	Output
0	0	A
0	1	B (shown)
1	0	C
1	1	D

1. A data selector which connects the output line, E, with the four input lines A, B, C and D.



2. A two-input data selector with only one control line.

numbers are sent straight through, while a logic 1 will make them cross over. As you can see, this concept of data selection can be very versatile.

The selector

The **selector** or **multiplexer** permits the selection of one of many inputs. For an understanding of how these work, we'll take as an example the 74LS253 multiplexer IC. This package houses two, four-input selectors and is known as a 'dual 4-line-to-1-line multiplexer'. It has a three state output which means that as well as offering the usual high and low outputs it can, under certain circumstances, present a high impedance output – this disconnects the IC from the system in which it is operating. Three state devices will be dealt with in greater detail later on.

The function table in *figure 3* illustrates the operating properties of this device. If N is the number of data inputs to be selected, then the number of select inputs needed by a multiplexer is n, where $2^n = N$. Thus, if there are sixteen inputs to be multiplexed, the number of select inputs is four ($2^4 = 16$).

Looking again at the function table for the 74LS253, we can see that there are two select inputs, four data inputs and one output. The symbols within the table mean the following: H/high; L/low; X/don't care; while Z corresponds to the high impedance output. As you can see, output Y is high only when certain input configurations are present:

3

Select Inputs		Data Inputs				Output Control	Output
B	A	C0	C1	C2	C3	G	Y
X	X	X	X	X	X	H	Z
L	L	L	X	X	X	L	L
L	L	H	X	X	X	L	H
L	H	X	L	X	X	L	L
L	H	X	H	X	X	L	H
H	L	X	X	L	X	L	L
H	L	X	X	H	X	L	H
H	H	X	X	X	L	L	L
H	H	X	X	X	H	L	H

3. Function table for the 74LS253 multiplexer IC.

B	A	C0	C1	C2	C3	Y
L	L	H	X	X	X	H
L	H	X	H	X	X	H
H	L	X	X	H	X	H
H	H	X	X	X	H	H

This can be summarised in the expression:

$$Y = \overline{A}\overline{B}C0 + \overline{A}BC1 + A\overline{B}C2 + ABC3$$

The arrangement of selector/multiplexer pins

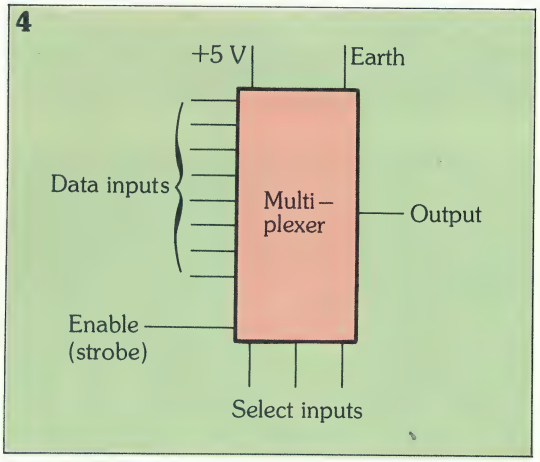
The block symbol in *figure 4* is a general representation of a multiplexer. The data input lines are on the left of the block, as is the **enabling input** or **strobe.** (As its name suggests, the enabling input *enables* the device to function.) The output is drawn to the right; the power supply and earth pin are at the top; and the select inputs are at the bottom of the diagram.

4. General representation of a multiplexer.

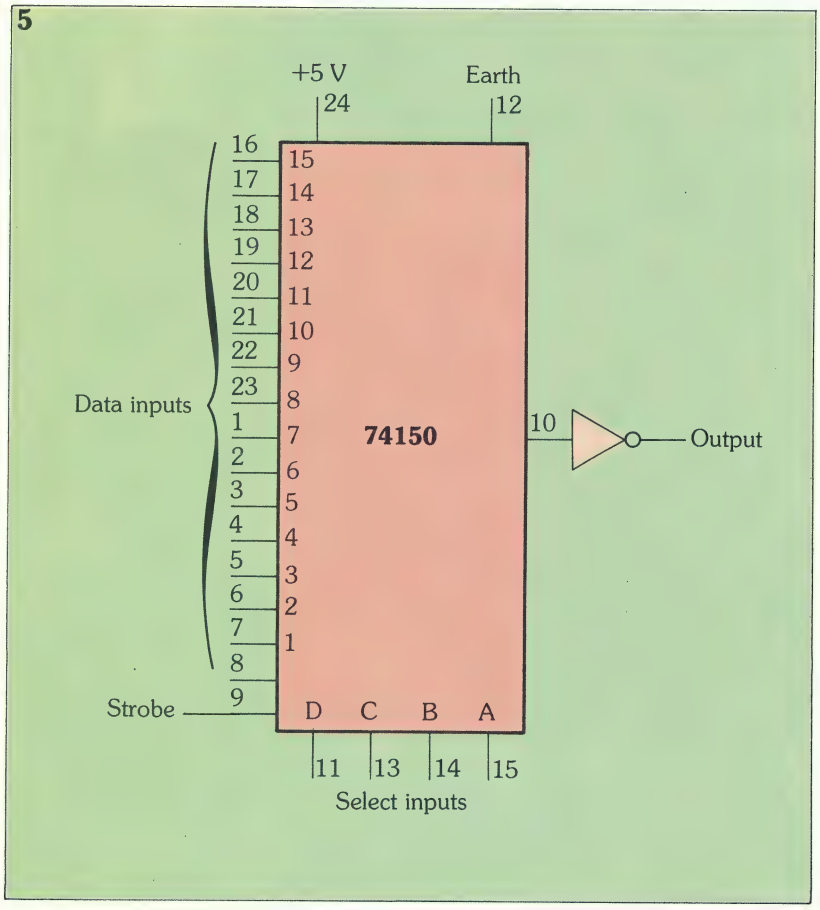
5. Pin configuration for the 74150 16-line-to-1-line data selector/ multiplexer.

6. Function table for the 74150 multiplexer.

4



5



The most widely used multiplexers in the TTL 7400 series are: the 74150 16-line-to-1-line selector/multiplexer; the 74151 8-line-to-1-line selector/multiplexer; and the 74153 dual 4-line-to-1-line selector/multiplexer. We shall look at each of these devices in turn.

74150 16-line-to-1-line data selector/multiplexer

The pin configuration (or connection diagram) for this IC is shown in *figure 5*. The output (pin 10) of this device inverts the data input, an inverter is therefore placed on this input, returning the data to its original value and producing a normal output. The function table for the 74150 is shown in *figure 6*.

6

Inputs					Output
Select				Strobe	
D	C	B	A	S	W
X	X	X	X	H	H
L	L	L	L	L	$\overline{E0}$
L	L	L	H	L	$\overline{E1}$
L	L	H	L	L	$\overline{E2}$
L	L	H	H	L	$\overline{E3}$
L	H	L	L	L	$\overline{E4}$
L	H	L	H	L	$\overline{E5}$
L	H	H	L	L	$\overline{E6}$
L	H	H	H	L	$\overline{E7}$
H	L	L	L	L	$\overline{E8}$
H	L	L	H	L	$\overline{E9}$
H	L	H	L	L	$\overline{E10}$
H	L	H	H	L	$\overline{E11}$
H	H	L	L	L	$\overline{E12}$
H	H	L	H	L	$\overline{E13}$
H	H	H	L	L	$\overline{E14}$
H	H	H	H	L	$\overline{E15}$

74151 8-line-to-1-line data selector/multiplexer

The connection diagram for the 74151 is shown in *figure 7:* the function table in *figure 8.* As this device has eight data inputs, there are three data selection inputs (remember $2^3 = 8$). This IC gives normal output Y, at pin 5 and an inverted output W, at pin 6. If the strobe input is set to logic 1, then the entire 74151 IC is disabled. Consequently output Y becomes 0 and output W 1, regardless of the state of the data and selection inputs. If the strobe input is set to logic 0, the device is considered to be undergoing a *strobe* and is enabled to function normally. A strobe

input can therefore be used to select one or more multiplexers from any number available: a strobe input therefore acts as a **chip selection** input.

74153 dual 4-line-to-1-line data selector/multiplexer

Figures 9 and *10* show the pin configuration and function table respectively for the 74153 multiplexer IC. The term 'dual' means that the package contains two multiplexers, both of which are built on the same chip. Although this does not mean that they can operate independently (there are only two selection inputs acting for both multiplexers), it is possible to select one of the multiplexers by applying the appropriate logic states at the strobe inputs. For example, when strobe 1G = 0 and strobe 2G = 1 we select multiplexer number 1. When strobe 1G =1 and strobe 2G = 0, the second multiplexer is chosen.

As you can see, multiplexer number 1 has its data inputs on pins 3,4,5 and 6, a strobe input on pin 1 and a single normal output on pin 7. The other multiplexer has its data inputs on pins 10,11,12 and 13, a strobe input on pin 15 and a single normal output on pin 9. The selection inputs are on pins 2 and 14.

A tree structure

If a multiplexer with more than eight or sixteen inputs is required, several devices must be linked together in what is known as a **tree structure.**

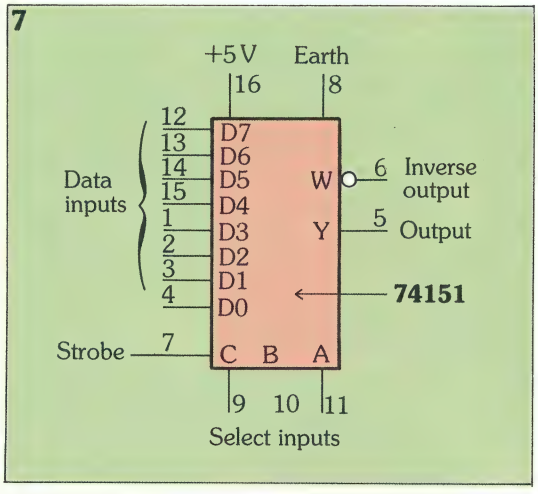
For example, a selector that multiplexes sixty four inputs can be made using nine, eight-input selectors connected as shown in *figure 11*. The eight selectors on the second level each choose one of eight inputs. The selector on the first level then chooses one of the 8 outputs of the second level selectors. Using this method, 64, 128 and 256-input selectors can be made.

CMOS selector/multiplexers

The CMOS 4000 series has a number of selector/multiplexers. The most commonly used CMOS selectors are the 4019 quad AND/OR selector IC, the 4512 eight-input multiplexer (with three state output), and the 4539 dual four-input multiplexer.

The pin configuration and function table for the 4019, for example, are shown in *figure 12*. This IC allows the inputs A and B to be selected by means of the control inputs S_A and S_B, which are active at a high level.

7



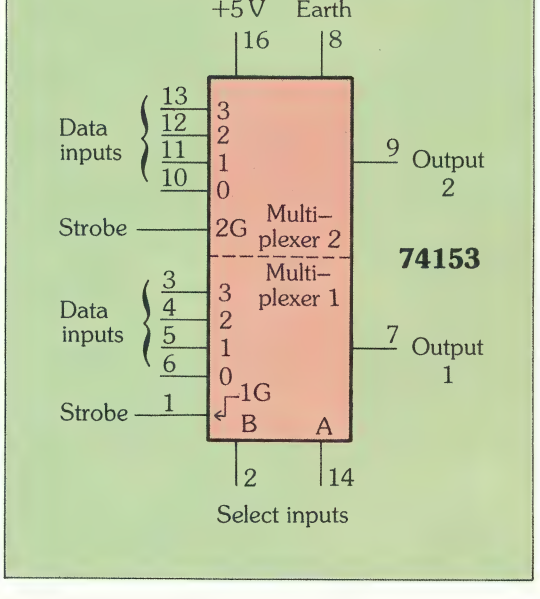
7. Pin configuration for the 74151 8-line-to-1-line data selector/multiplexer.

8

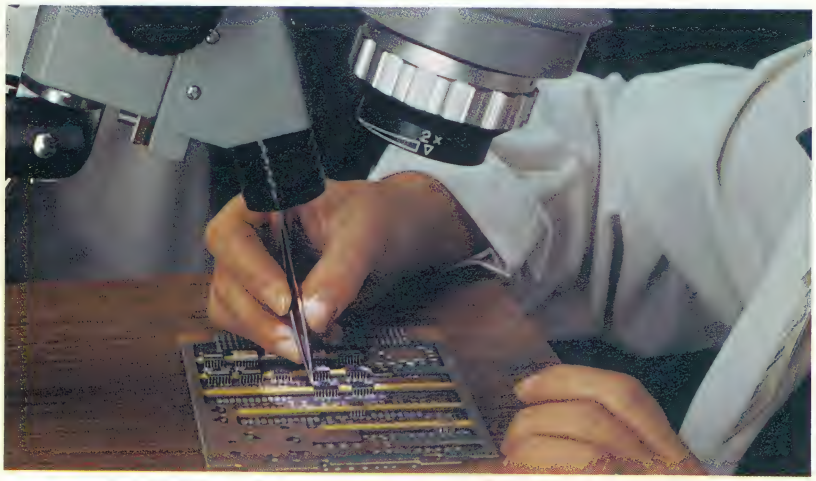
Inputs				Outputs	
Select			Strobe		
C	B	A	S	Y	W
X	X	X	H	L	H
L	L	L	L	D0	$\overline{D0}$
L	L	H	L	D1	$\overline{D1}$
L	H	L	L	D2	$\overline{D2}$
L	H	H	L	D3	$\overline{D3}$
H	L	L	L	D4	$\overline{D4}$
H	L	H	L	D5	$\overline{D5}$
H	H	L	L	D6	$\overline{D6}$
H	H	H	L	D7	$\overline{D7}$

8. Function table for the 74151 multiplexer.

9



9. Pin configuration for the 74153 dual 4-line-to-1-line data selector/multiplexer.



Ace Photo Agency/V.S. Studios

Above: examination of a circuit board under a microscope.

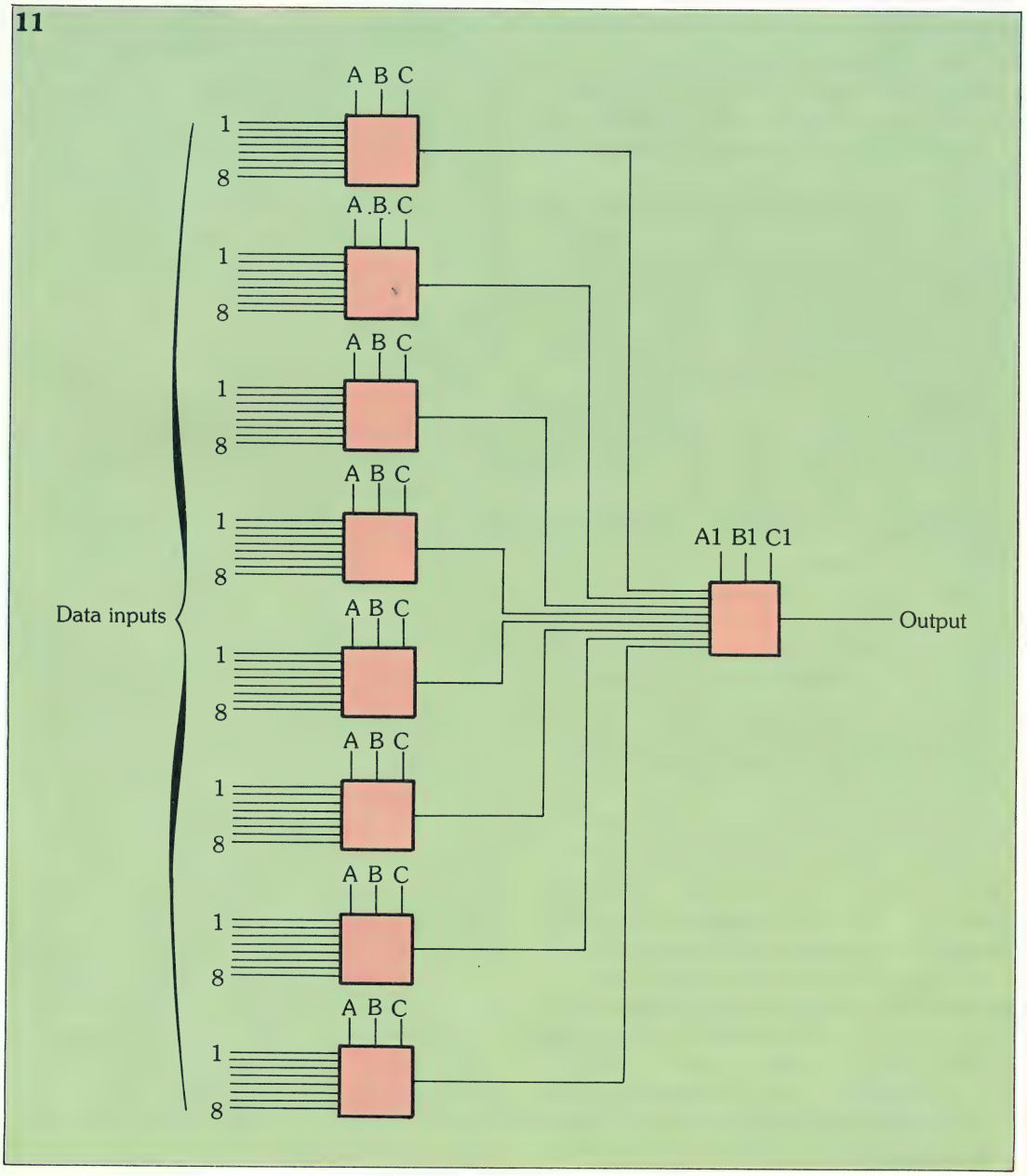
10. Function table for the 74153 multiplexer.

10

Select Inputs		Data Inputs				Strobe	Output
B	A	C0	C1	C2	C3	G	Y
X	X	X	X	X	X	H	L
L	L	L	X	X	X	L	L
L	L	H	X	X	X	L	H
L	H	X	L	X	X	L	L
L	H	X	H	X	X	L	H
H	L	X	X	L	X	L	L
H	L	X	X	H	X	L	H
H	H	X	X	X	L	L	L
H	H	X	X	X	H	L	H

11. An example of a tree structure. Nine eight-input selectors connected to multiplex 64 inputs.

11



Demultiplexing an input among several outputs

A data routing unit known as a **demultiplexer** is shown in *figure 13*. This device operates in the opposite way to a data selector, that is to say, it routes data coming from a single input, to one of many outputs. The demultiplexer in *figure 13* routes data from the single input F, to any of the outputs G,H,J or K.

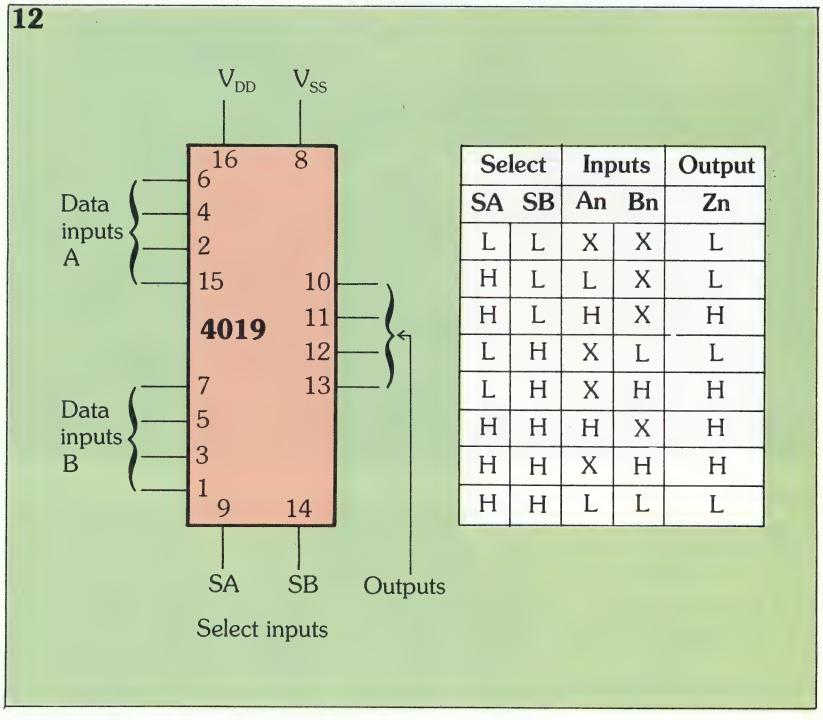
Demultiplexers

The demultiplexer shown by the block diagram in *figure 14* allows one of eight outputs to be selected and connected to the input. Selection is controlled by applying a 3-bit word to the selection inputs. (Remember, a 3-bit word is used because $2^3 = 8$, which is the number of outputs available.)

There are many demultiplexers in the TTL 74 series, and we shall look at the 74154 IC. The schematic diagram of the 74154 decoder/demultiplexer is shown in *figure 15*. The G2 strobe pin is connected to earth, while the G1 strobe input is used as an input. Suppose that the selection inputs DCBA are set to select output 6. You can see from the function table (*figure 16*) that when G1 is H, i.e. logic 1, output 6 is 1, and when G1 is L, i.e. logic 0, output 6 is 0. All the other outputs will be at logic 1. It is evident that this IC can be used to demultiplex data from one input to one of sixteen different output lines. Data is transferred directly along the output lines without inversion, as you can see in *figure 16*.

A demultiplexer function can be performed by the following TTL devices: the 74154 4-line-to-16-line decoder/demultiplexer, the 74159 4-line-to-16-line decoder/demultiplexer with open collector outputs, and the 74155 dual 2-line-to-4-line decoder/demultiplexer.

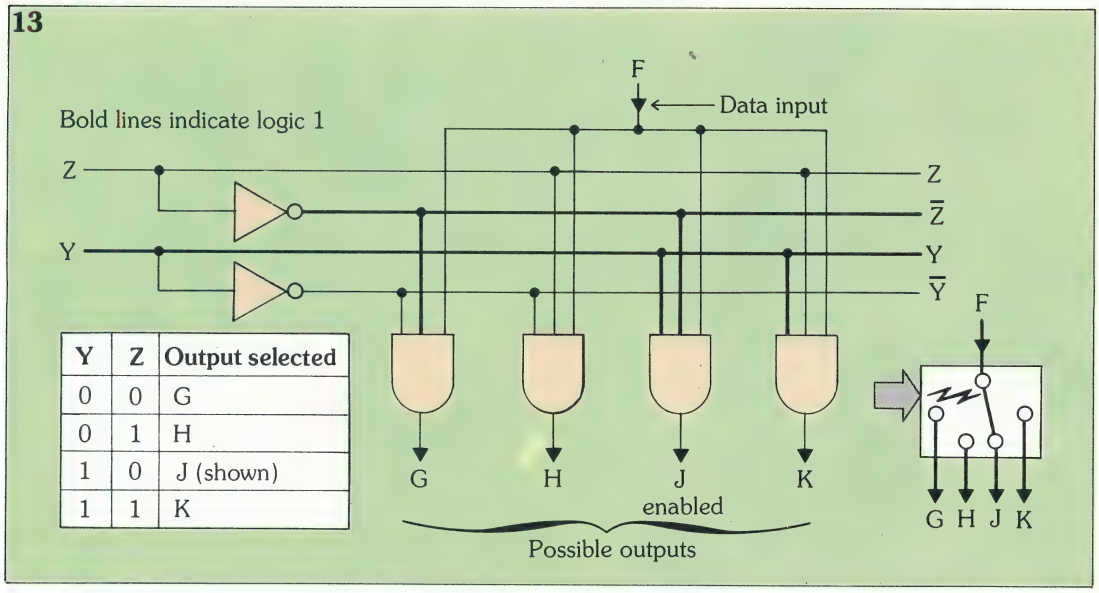
12



Select		Inputs		Output
SA	SB	An	Bn	Zn
L	L	X	X	L
H	L	L	X	L
H	L	H	X	H
L	H	X	L	L
L	H	X	H	H
H	H	H	X	H
H	H	X	H	H
H	H	L	L	L

12. Pin configuration and function table for the 4019 quad AND/OR Selector IC.

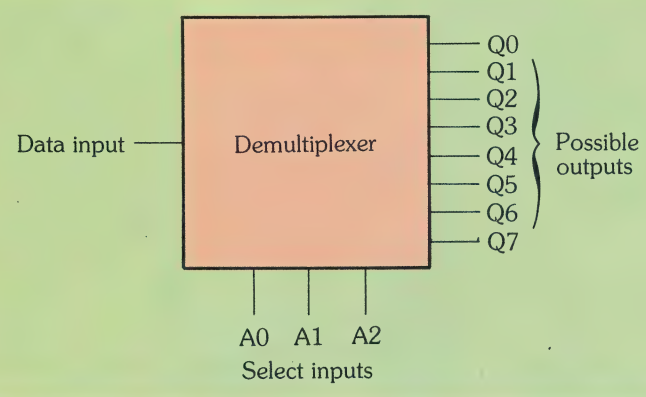
13



Y	Z	Output selected
0	0	G
0	1	H
1	0	J (shown)
1	1	K

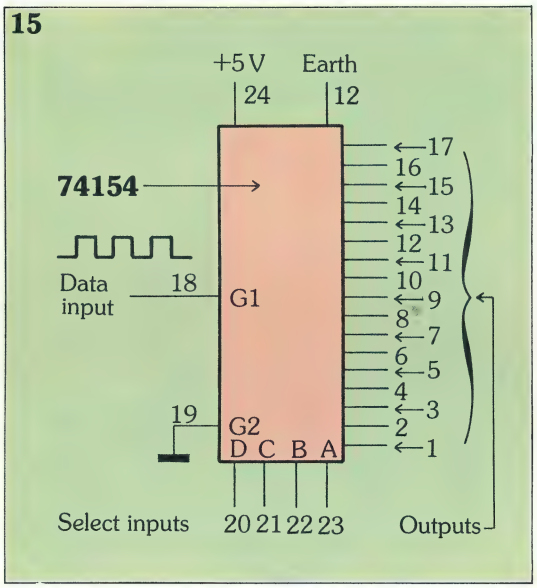
13. Block diagram of a demultiplexer which routes data from a single input, F, to one of many outputs G, H, J or K.

14



14. Block diagram of demultiplexer allowing one of eight outputs to be selected and connected to the input.

15



15. Schematic diagram of the 74154 decoder/demultiplexer.

16. Function table for the 74154 decoder demultiplexer.

17. A 4-bit binary adder.

16

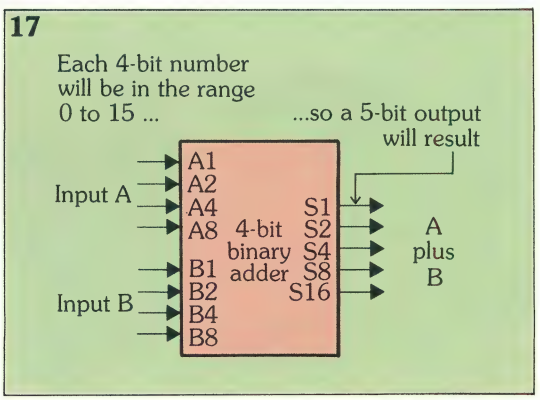
Inputs						Outputs															
G1	G2	D	C	B	A	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
L	L	L	L	L	L	L	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
L	L	L	L	L	H	H	L	H	H	H	H	H	H	H	H	H	H	H	H	H	H
L	L	L	L	H	L	H	H	L	H	H	H	H	H	H	H	H	H	H	H	H	H
L	L	L	L	H	H	H	H	H	L	H	H	H	H	H	H	H	H	H	H	H	H
L	L	L	H	L	L	H	H	H	H	L	H	H	H	H	H	H	H	H	H	H	H
L	L	L	H	L	H	H	H	H	H	H	L	H	H	H	H	H	H	H	H	H	H
L	L	L	H	H	L	H	H	H	H	H	H	L	H	H	H	H	H	H	H	H	H
L	L	L	H	H	H	H	H	H	H	H	H	H	L	H	H	H	H	H	H	H	H
L	L	H	L	L	L	H	H	H	H	H	H	H	H	L	H	H	H	H	H	H	H
L	L	H	L	L	H	H	H	H	H	H	H	H	H	H	L	H	H	H	H	H	H
L	L	H	L	H	L	H	H	H	H	H	H	H	H	H	H	L	H	H	H	H	H
L	L	H	L	H	H	H	H	H	H	H	H	H	H	H	H	H	L	H	H	H	H
L	L	H	H	L	L	H	H	H	H	H	H	H	H	H	H	H	H	L	H	H	H
L	L	H	H	L	H	H	H	H	H	H	H	H	H	H	H	H	H	H	L	H	H
L	L	H	H	H	L	H	H	H	H	H	H	H	H	H	H	H	H	H	H	L	H
L	L	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	L
L	H	X	X	X	X	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
H	L	X	X	X	X	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
H	H	X	X	X	X	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H

Structural units which add numbers

The combinational building blocks that perform arithmetical functions in digital systems are known as **adders**. A 4-bit binary adder is shown in *figure 17*. As you can see, it takes two, 4-bit binary numbers, A and B, and produces a 5-bit binary sum of A plus B. (Not the Boolean OR function this time, but an arithmetic sum.)

The label 'S16' tells us that the fifth output wire has a value of 16, compared with the values of 8, 4, 2 and 1 for the other four output wires. As each input number can run from zero (0000) to fifteen (1111), the sum can run from zero (00000) to thirty (11110). If the sum output had only four wires, like the inputs, it could only go up to fifteen.

17



This building block could be built using separate gates, taking the sum-of-products approach right from the truth table, and then trying to simplify it. However, the sum-of-products network would need 256 (2^8) eight-input AND gates to handle all the input combinations possible in eight wires. The output OR gates would have to handle as many as 128 inputs. This would be virtually impossible to build, as each output OR function would have to be processed in two or more stages – for example with sixteen eight-input OR gates which would feed a seventeenth. It is therefore more sensible to adopt a different approach.

Binary adder design

Figure 18 holds the solution to the problem of adder design. Here we see the numbers

44 and 58 added together to get the sum of 102 – both in the customary decimal way (*figure 18a*) and in the binary way (*figure 18b*). Remember, we're talking about *pure* binary numbers here, not binary coded decimal numbers. We shall see how pure binary adders are used to add binary coded decimal numbers in a later chapter.

Binary numbers are added in the same way as decimal numbers: the two numbers are aligned so that the *ones* column of the first number lies above the *ones* column of the second number, and so on for all the number positions. Then, beginning on the right (the *least significant* end), add one pair of digits at a time. This is a binary sum and so it must be remembered that the symbols 0 and 1 provide the following equations:

```
  0      0      1
 0+     1+     0+
 ---    ---    ---
  0      1      1
```

When there are two *ones* in the same column we get the following result:

```
   1
  1+
  --
  10
```

If you recall, 10 in the binary system corresponds to a decimal 2, so this is the correct answer. In this case then, the extra 1 in the answer is **carried** and added to the next column on the left.

For example, adding the numbers 110 and 111 gives us:

```
   110
  111+
  ----
  1101
   11   carry bits
```

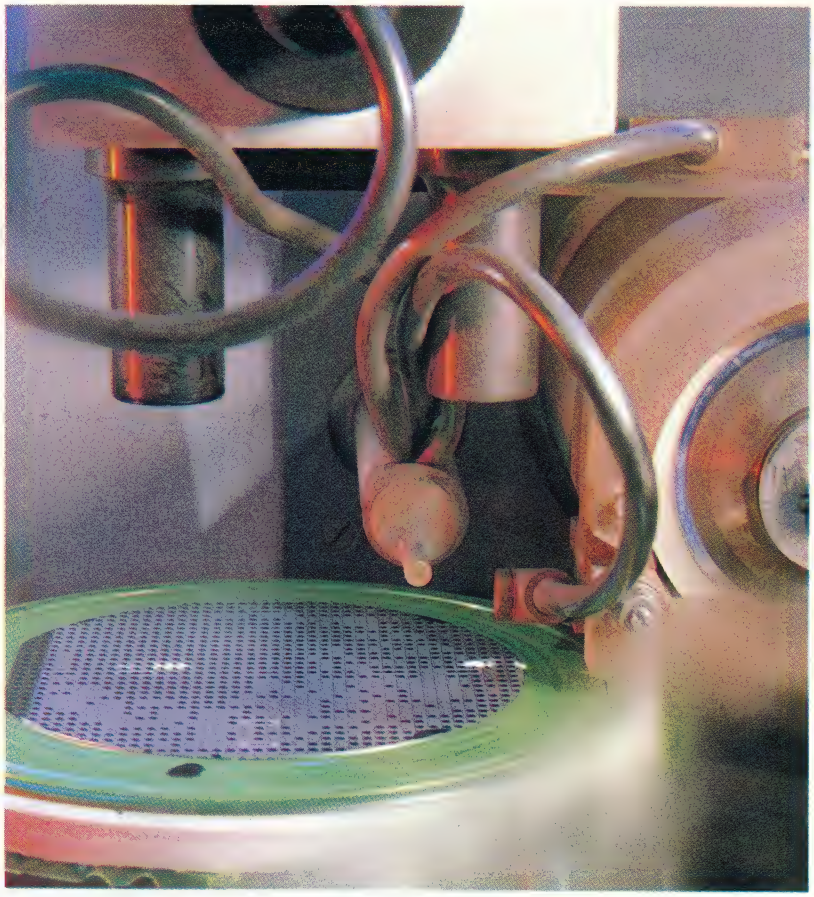
If you follow the addition through, you will see where the carry bits come from, and how they operate.

Now do you see a simpler way to handle binary addition? To add two 4-bit numbers, all we need is four 1-bit adders. That way, we can handle this complex combinational circuit in four easy stages. The key to digital system design is to break big tasks down into small tasks.

One's and two's complements

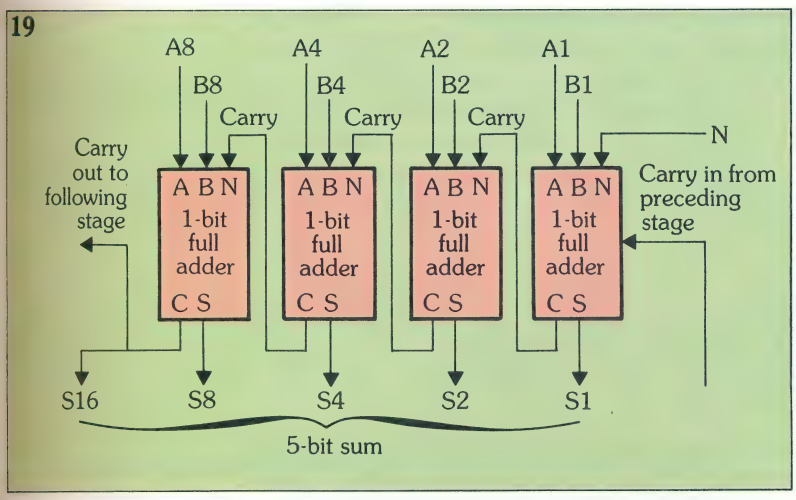
The **one's** and **two's complements** of binary numbers are very useful in many applications and operations. The one's complement of a binary number is simply obtained by inverting each bit, i.e. 1 is changed to 0 and vice versa. For example, the one's complement of 1011 is 0100. When the one's complement of a number is added to the original number, a 1 will be the result in each column.

The two's complement of a binary number is obtained by taking the one's complement and adding 1 to the least



Above: mechanical separation of individual chips from the wafer. (Photo: ITT).

18. (a) Decimal and **(b)** binary addition.

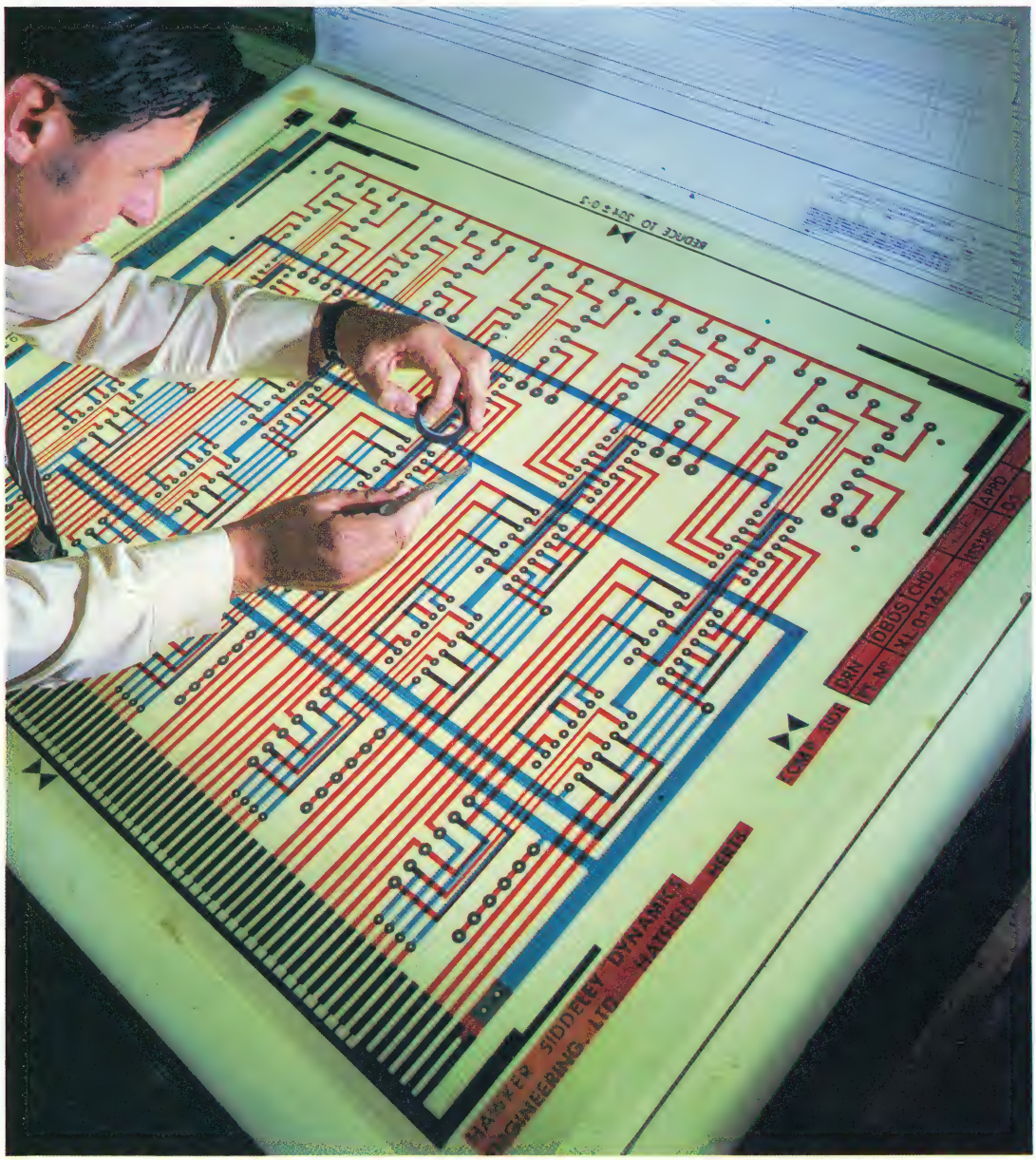


19. Typical arrangement of a 4-bit adder.

significant digit. To find the two's complement of the number 10101, for example, the one's complement is calculated first. This is 01010. Then 1 is added, giving the result:

$$\begin{array}{r} 01010 \\ \quad\quad 1+ \\ \hline 01011 \end{array}$$

As you may have gathered, computers and calculators rely on adder circuits to perform all their arithmetical functions. Multiplication is in effect repeated addition; division is repeated subtraction; subtraction itself is carried out by adding a negative number to a positive number. One of the main applications of one's and



Right: a printed circuit board being designed.

Ace Photo Agency/John Price

two's complements is in the representation of negative numbers – converting them to a form that can be added in the normal manner. Before we look at this in more detail, however, the binary representation of positive and negative numbers needs to be considered.

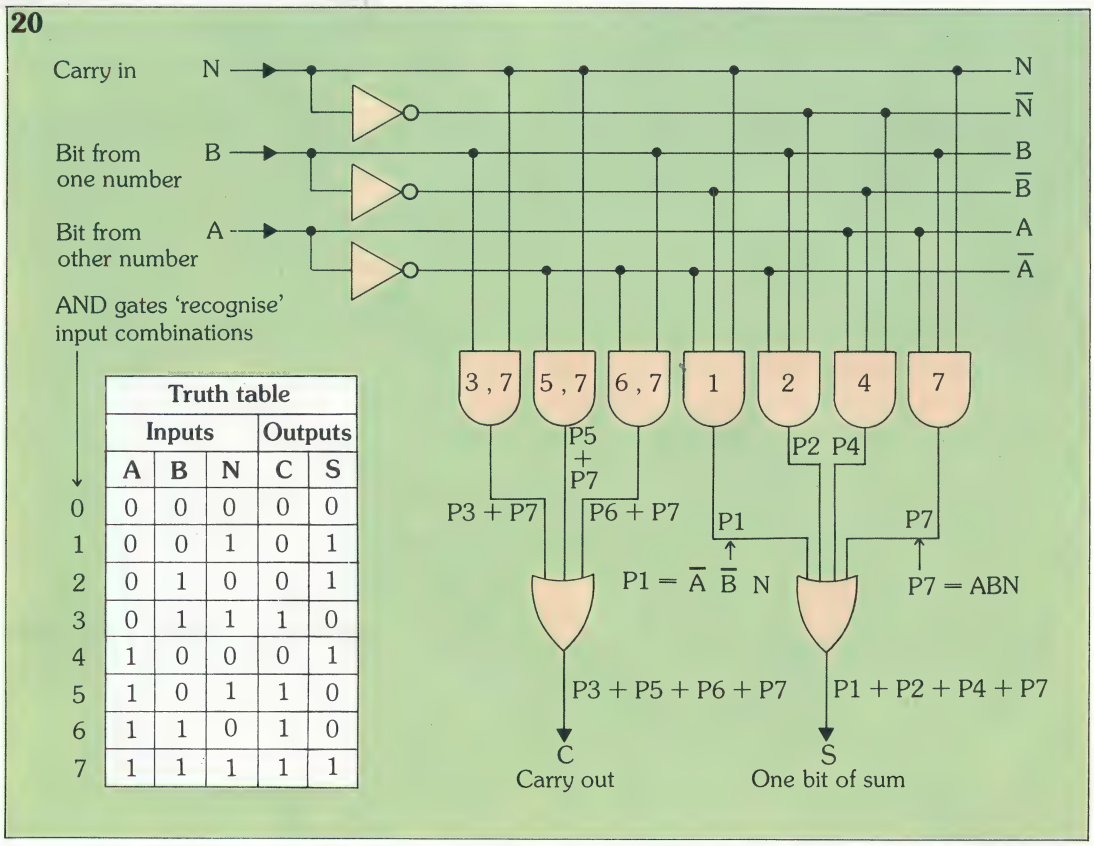
Generally speaking, the decimal numbering system uses the symbols + and – to indicate a positive or negative number. As computers only use bits to represent numbers these signs cannot be employed, so another way of indicating positive and negative numbers is needed.

The method employed first makes all numbers up to the same number of bits by adding zeros to the left. For example, using a three-bit number length, 5 = 101 but 3 = 011. This number is preceded by a special bit which indicates its sign (i.e. + or –). Positive numbers are preceded by a 0, separated from other digits by a point, e.g. +5 = 0.101. Negative numbers are represented by taking the two's complement of the positive number, e.g. –5 = 1.011. The 1 before the point indicates that the number is negative, and to find its value, the two's complement of both sides of the decimal must be calculated.

Let's consider the addition of a positive and a negative number: 0.011 (+3) and 1.011 (–5). Adding them together:

0.011
1.011+
1.110

The result is a negative number (the sign digit is 1); its value is found by taking the two's complement, which is 0.010, i.e. 2. The result of the calculation is therefore –2.



	Truth table				
	Inputs			Outputs	
	A	B	N	C	S
0	0	0	0	0	0
1	0	0	1	0	1
2	0	1	0	0	1
3	0	1	1	1	0
4	1	0	0	0	1
5	1	0	1	1	0
6	1	1	0	1	0
7	1	1	1	1	1

20. Truth table for a 1-bit full-adder.

Let's now take an example of subtraction where a negative number, say, –5 (1.011) is subtracted from a positive number, say, 2 (0.010). The adder circuit used to perform the calculation cannot subtract; so the two's complement of the negative number is added to the positive number. The original calculation:

0.010
1.011–

therefore becomes:

0.010
0.101+
0.111

The result, 7, is a positive number as the sign digit is 0.

How 1-bit adders can add multibit numbers

The typical arrangement of many 4-bit adders is shown in *figure 19*. The 7483 is one such device, and is made up of four identical 1-bit adders. Each of these adders, adds two bits, A and B, and takes in a third carry bit. Each 1-bit adder produces a sum, S, as output and a carry bit to pass to the next stage. Naturally, any number of adders can be connected in sequence, so that numbers with any amount of bits can be processed.

However, for the 4-bit adder shown in *figure 17*, we need only four 1-bit adders connected as shown in *figure 19*. This is because there are no preceding or following stages of addition – a 0 is supplied to the carry input, N, of the first (least significant) adder.

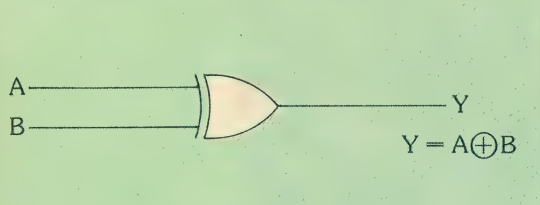
How a 1-bit full adder works

The truth table for a 1-bit full adder is given in *figure 20*: it shows the possible combinations for the inputs A,B and N, and the outputs C and S. The network shown here is probably the simplest way of building this circuit, if only AND, OR and NOT gates are used.

Every combination of the truth table is characterised by a number from 0 to 7 – which corresponds to the binary number obtained from this combination. The AND function of a given input combination is indicated by a P (for product) which has the same number subscripted. So, for example:

$P_1 = \bar{A}\bar{B}N$ and $P_7 = ABN$

21



A	B	Y
0	0	0
0	1	1
1	0	1
1	1	0

21. The logic symbol and the truth table for the exclusive OR gate.

Binary arithmetical circuits

The operation of addition, like that of subtraction, multiplication and division is carried out in a special part of the computer known as the arithmetic and logic unit (**ALU**). We shall now look at the combinational units that perform these functions.

The exclusive OR gate

We have looked at the fundamental AND, OR, NAND and NOR logic gates in earlier chapters. The **exclusive OR (XOR)** gate must now be introduced as it has particular importance in arithmetical applications. The exclusive OR operation is represented by the symbol ⊕, and is characterised by the logical expression:

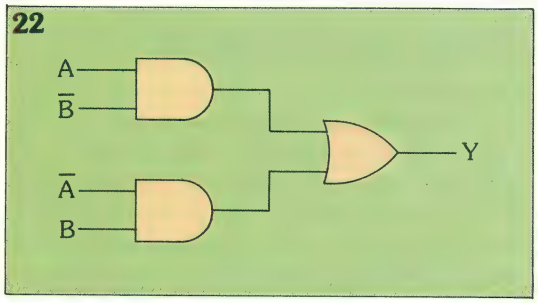
$$Y = A \oplus B$$
$$= \overline{(A + \bar{B})\,(\bar{A} + B)}$$
$$= A\bar{B} + \bar{A}B$$

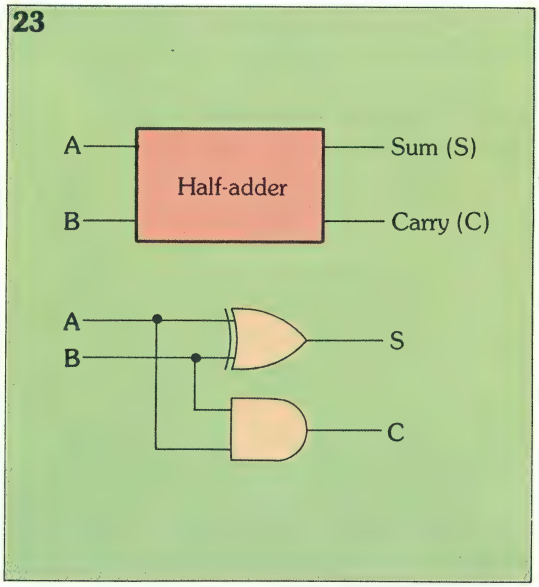
The truth table and the logic symbol for this function are shown in *figure 21*. As you can see, the exclusive OR function, when applied to two variables only, is only 1 if either one or the other – but not both – of the input variables is 1.

Before we look at the integrated devices that perform the exclusive OR function, we shall see how this function can be made by combining other logic gates. Look at the truth table and the logical expression for the exclusive OR function. You can see that the output Y is given by the sum of two previous products – so this circuit can be made from two AND gates on the second level and an OR gate on the first level (*figure 22*).

The exclusive OR gate is also known as a **modulo 2 adder** and a **half-adder**. From what has been said, we can see that the logic function which carries out the 1-bit sum of the two numbers, A and B, is none other than the XOR function, while the carry C, is due to the A AND B operation:

Addend A	Augend B	Sum S	Carry C
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1





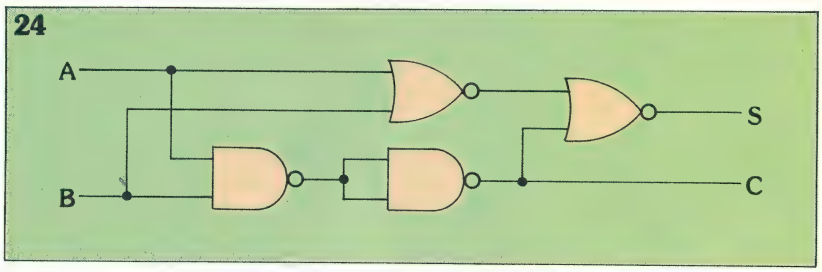
22. An exclusive OR gate using AND/OR logic.

23. A half-adder that uses an exclusive OR gate and an AND gate to generate the carry.

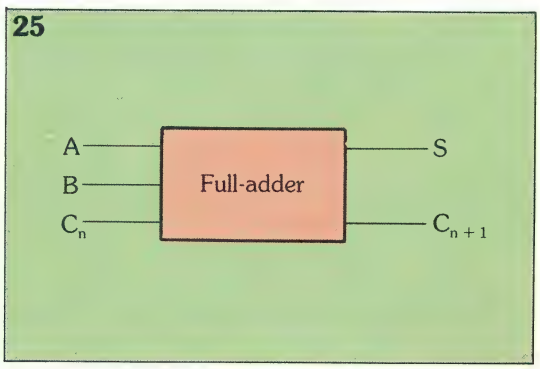
The limitation of this 1-bit addition is that the carry from any previous addition is not considered – which is why this is known as a half-adder. The logic symbol for a half-adder is shown in *figure 23* and in practice it is made from an XOR and an AND gate. If the circuit has to be built with conventional NAND and NOR gates then the scheme illustrated in *figure 24* is used. This employs the standard TTL 7400 and 7402 ICs.

However, as addition must usually be done with a carry input, a full-adder must be used. The logic symbol for this is shown in *figure 25*. This unit has three inputs: A,B and C_n, and two outputs, S and C_{n+1}. Input C_n represents the carry from the preceding bit, while output C_{n+1} is the carry for the following bit. The table in *figure 26* summarises the whole operation.

Full-adders can be made in two different ways – by using two half-adders and an OR gate, or by using AND, OR and NOT gates. *Figure 27* shows a full-adder made from two half-adders and an OR gate. The method using AND, OR and NOT gates was, of course, shown earlier in *figure 24*.



24. Conventional NAND and NOR gates used to build the half-adder circuit of *figure 23*.



25. Logic symbol for a full-adder.

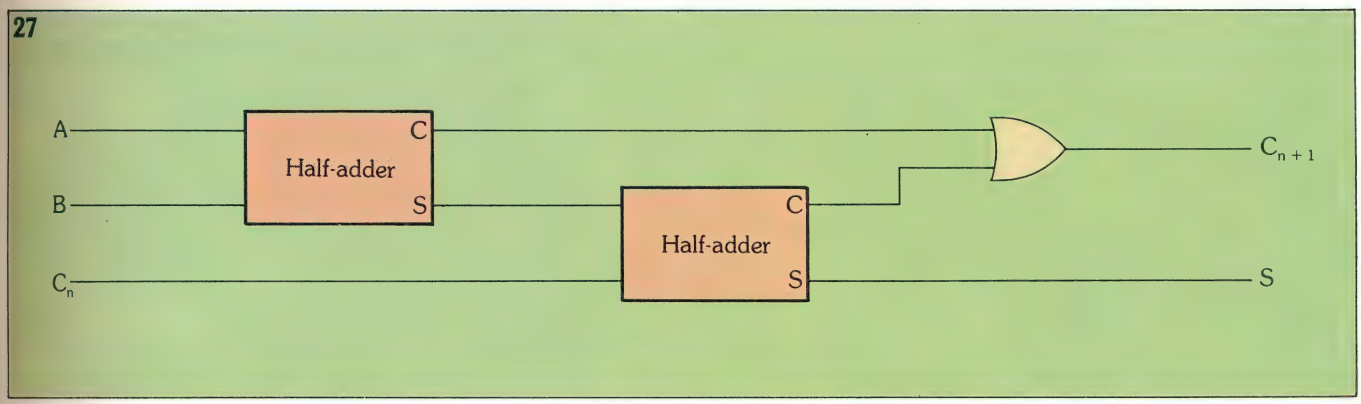
Parallel binary addition

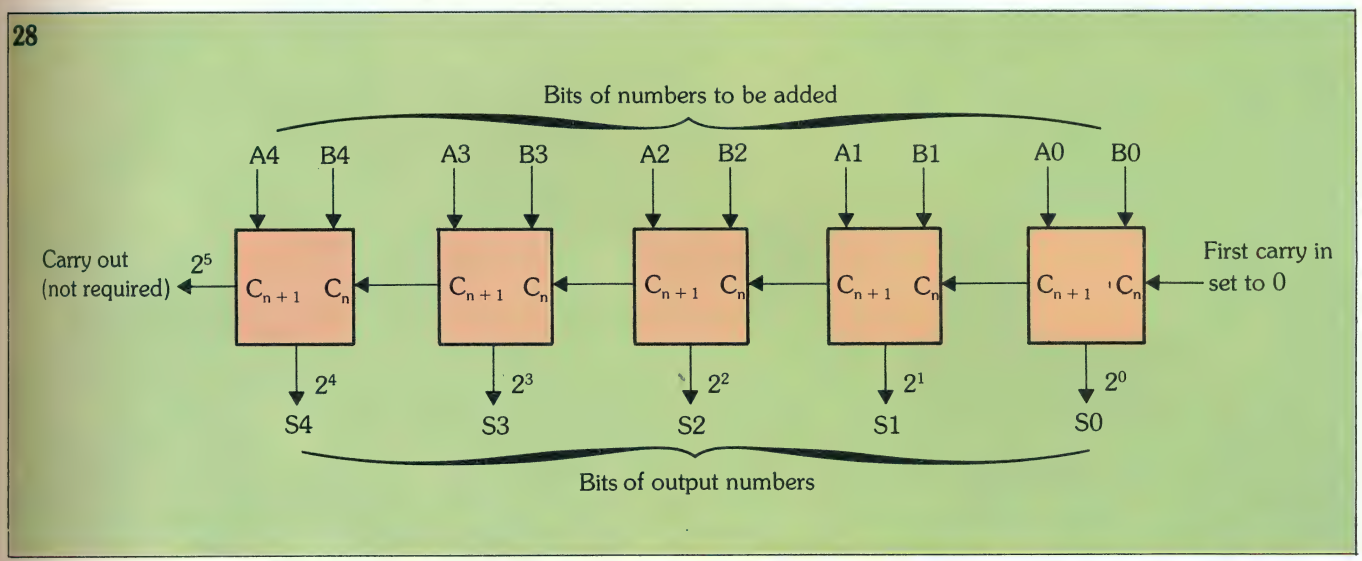
To add two binary numbers, the same number of adders as there are bits in the two numbers must be used. For example, when adding two, 5-bit numbers, a carry input is not needed in adding the least significant digits, so a half-adder or a full-adder with the carry input set to 0 is used. The remaining addition modules, on the other hand, *must* be full-adders, as the sum of two bits can involve carrying. The complete arrangement is shown in *figure 28*. This is a **cascade** circuit, and the propagation delay of the units must be taken into account, as it is crucial in many applications.

26

Addition			Output	
C_n	A	B	S	C_{n+1}
0	0	0	0	0
0	1	0	1	0
0	0	1	1	0
0	1	1	0	1
1	0	0	1	0
1	1	0	0	1
1	0	1	0	1
1	1	1	1	1

26. Truth table for a full-adder.



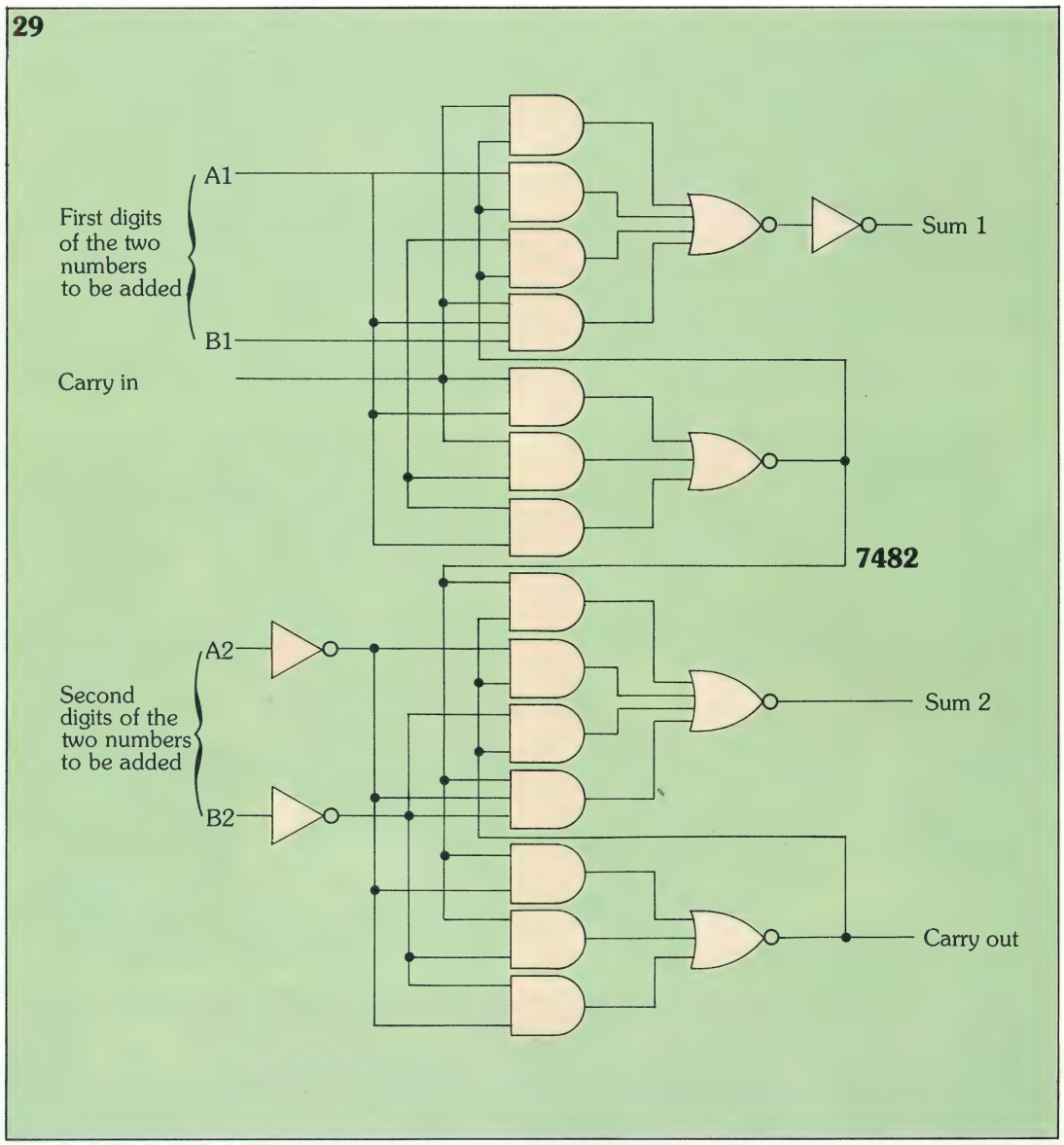


27. A full-adder made from two half-adders and OR gate.

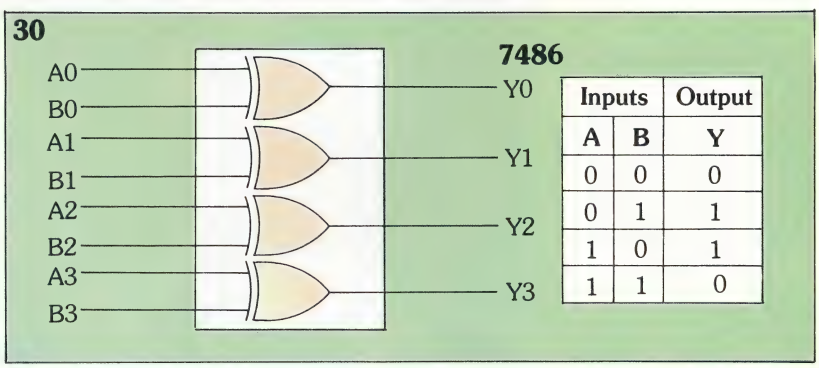
28. Adding two 5-bit numbers – producing a cascade circuit.

Right: Infra-red map of the Lake Erie/Lake Ontario area produced from data supplied by Applications Explorer Mission 1. The data was sent in real time to the NASA Flight Centre.

The Research House/NASA



29. Circuit diagram for the 7482 2-bit adder.



Inputs		Output
A	B	Y
0	0	0
0	1	1
1	0	1
1	1	0

30. Functional diagram and truth table for the 7486 IC.

Arithmetic devices in the TTL series

There are many arithmetic devices in the TTL 74 series allowing different methods of arithmetical manipulation to be carried out.

The 7480 is a single bit gated full-adder, with complementary inputs, and can give a normal and a complementary output. This circuit actually uses DTL logic for the inputs, while the sum and carry outputs are conventional TTL with a high fan-out.

The 7482 contains two 2-bit full-adders, and its circuit arrangement and connection details are shown in *figure 29*. Each full-adder of the IC performs the parallel addition of two 2-bit numbers.

Finally, consider the 7486, this IC contains four, two-input XOR gates. The block diagram of the inputs, outputs and the truth table for this device are shown in *figure 30*.

A summary of combinational circuits

So far we have looked at the main building blocks that make up digital systems: various types of code converters, data transfer units and adder units. All of these decision units have one important characteristic in common, that is, for each combination of signals received in input, there is a certain predetermined combination of output signals. These input and output combinations can be summarised in an orderly way in a truth table.

Boolean algebra provides a shorthand method of describing the logic that allows you to establish which signals are produced in a circuit's output. We have seen that a circuit which is designed from a truth table is not necessarily the most economical one, and can probably be simplified by taking 'undesirable' impossible truth table conditions into account. Some gates, different to those originally used in a circuit, can be employed to make better use of the ICs available, and we also know how a combinational design problem can be dealt with in different stages to simplify the circuit.

We have therefore improved our knowledge of the principles of digital electronics, not only theoretically, but from the practical point of view. This basic knowledge will be of great assistance to us later on, particularly when we look at sequential networks and more complex digital systems.

Glossary

ALU	arithmetic and logic unit. The part of a computer's central processor that handles arithmetic and logic operations
demultiplexer	digital device used to route data coming from one input line, to one of many output lines
exclusive OR	known as a XOR gate and represented by the symbol ⊕. This gate gives an output of logic 1, when one or the other, but *not both* of its inputs is 1
full-adder	circuit that adds two 1-bit binary numbers, but cannot accept a carry input
multiplexer	digital device used to route data from one of many input lines to one output line
strobe	enabling input that controls the operation of a multiplexing or demultiplexing circuit. When the circuit is switched on it is said to be *undergoing a strobe*
tree structure	method of connecting selector circuits to enable them to multiplex or demultiplex a great number of input or output lines

11

SOLID STATE ELECTRONICS

Operating properties of transistors

Operating properties

As we now know, semiconductor components are capable of performing many different functions. For example: diodes can be used either to rectify, clamp or regulate voltages or to limit waveforms; transistors can be used to either switch or amplify electrical signals.

As there are many different types of transistor or diode, circuit designers are able to choose the most suitable for their particular needs. Semiconductor manufacturers specify the properties, or **characteristics,** of every component in a standard form, normally tables. These tables give the value of each parameter under specific test conditions. Alternatively, component characteristics can be presented as a set of graphs, showing how the parameters vary over the whole range of operating conditions.

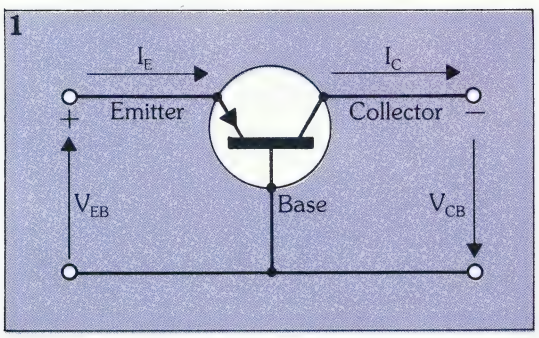
In addition to the description of a component's properties, the limits of the operating conditions must also be known as the component will be damaged if operated beyond these limits. The maximum permitted values of the collector voltage, power dissipation and the junction temperature of a transistor, for example, will be given for every operating condition.

The performance of a semiconductor component is reflected in its **electrical** characteristics. However, since some parameters vary over a range of temperatures, circuit designers must also have information regarding **thermal** characteristics. Electrical characteristics are usually measured at a free air temperature of 25°C; thermal characteristics indicate the way in which these values vary with increases or decreases in temperature.

Electrical characteristics

We already know that the overall performance of an electrical device can be defined by the relationship between the voltage applied to the device and the current flowing through it. This is reasonably easy to find out for a diode as there are only two terminals: therefore only one current needs to be measured for each applied voltage.

A transistor, on the other hand, has three terminals. To obtain a full description of its characteristics *two* voltages and *two* currents need to be measured for *every* state; more than one graph is therefore required to show the information needed. A transistor connected in common base mode is shown in *figure 1*: the currents and voltages used to indicate its electrical characteristics are marked. In this example, the emitter-base voltage V_{EB} and the emitter current I_E are taken as input values: the collector-base voltage V_{CB} and the collector current I_C are measured as outputs.



1. Transistor connected in common base mode.

The polarity of the voltages is taken with reference to the common terminal – in this case the base. A current is considered positive if the electrons flow *out* of the transistor and negative if they flow *into* it. The input characteristics of the transistor can be shown by an I_E *vs* V_{EB} curve, and the output characteristics by an I_C *vs* V_{CB} curve. However, when analysing these curves you must remember that the collector current is dependent upon the emitter current. To observe the effect on the

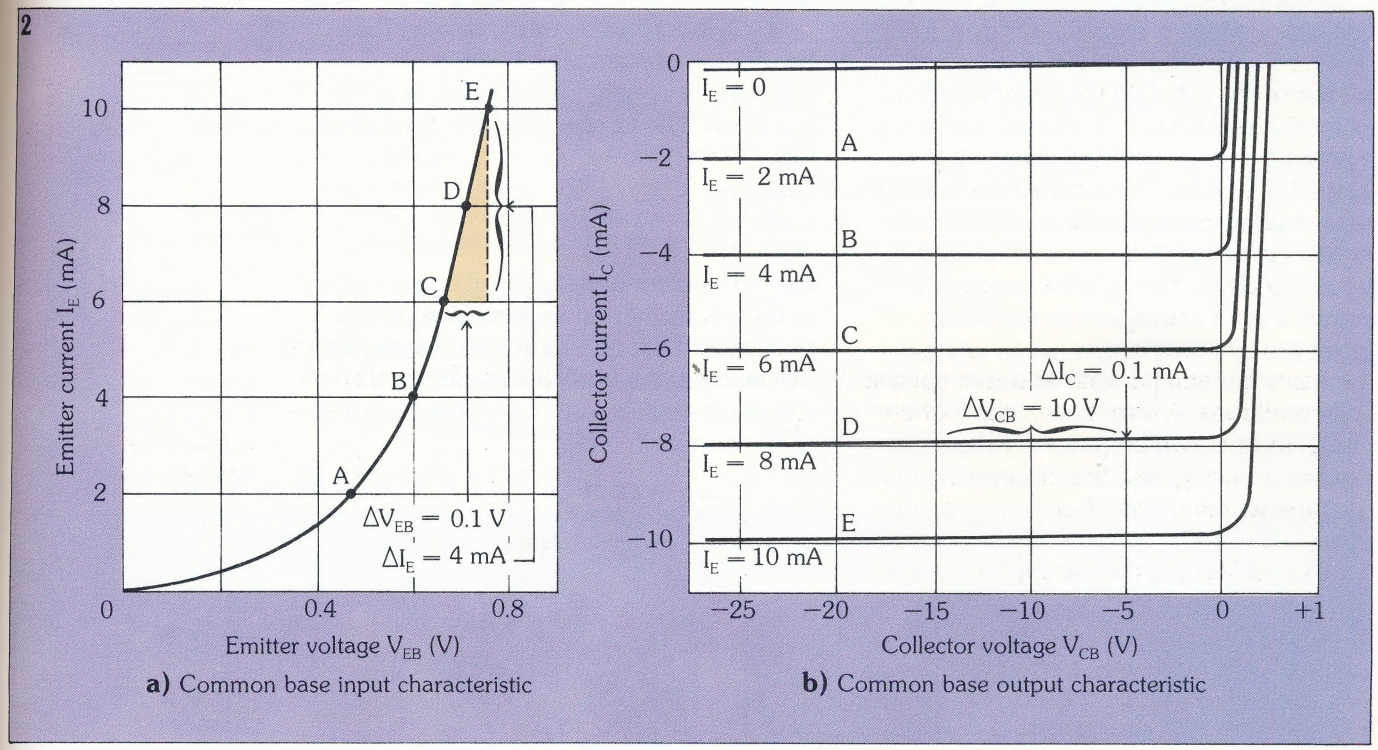
collector current that varying the collector voltage will have, the emitter current must be kept constant for each curve. The resulting graph thus shows the output characteristic for a particular value of emitter current.

Therefore, to obtain a complete range of information about a transistor, it is necessary to draw different curves for different emitter currents. In this way the result is a complete family of curves – it is useful to see how this family of output characteristics is obtained.

simple reverse polarized junction. The current flowing will be the reverse saturation current (I_{CBO}) and will remain substantially constant as the voltage increases. This is the current marked $I_E = 0$ in *figure 2b*.

If the emitter current is now increased to 1 mA (*figure 2a* shows this as being point A) and if this is kept constant while the collector voltage is varied, the new collector curve will be that marked A in *figure 2b*. The emitter current will pass through the collector no matter what the collector-base voltage level.

2. Characteristics for the p-n-p common base arrangement: **(a)** emitter; (b) collector.



a) Common base input characteristic

b) Common base output characteristic

Common base characteristics

If you consider the p-n-p common base arrangement in *figure 1*, then the possible characteristic of the emitter junction is shown in *figure 2a*. The particular values indicated in this and the following graphs are purely hypothetical: reference should be made to the manufacturer's reference data for any chosen transistor. This graph shows how the emitter current I_E increases with the rise in the emitter biasing voltage V_{EB}. The curve is called the **emitter characteristic,** or sometimes the **input characteristic.**

Next let's look at the collector circuit. When the emitter current is zero, the collector curve will be similar to that of a

If the emitter current is further increased, then curves B,C,D and E of *figure 2b* illustrate the characteristics for I_E, at 2 mA, 3 mA, 4 mA and 5 mA respectively. The precise increase in the collector current will be a $\propto$ multiplied by the increase in the emitter current. Remember, $\propto$ is the common base current transfer ratio which, in this case, is 0.975. Since the emitter current has been increased by 1 mA, the curves will each be 0.975 mA apart. All of these curves make up a family of collector characteristics. These output characteristics clearly show the effect of the emitter current and the collector voltage on the collector current, and are probably the most useful of the transistor characteristics.

The emitter characteristic that we have seen (*figure 2a*) has been represented by a single curve. However, variations in the collector voltage do have a limited effect on the emitter characteristic – to be correct, a family of emitter curves should be drawn, each curve being taken at a constant collector voltage. In reality though, the effect of the collector voltage is of such minor importance that it can be ignored in the majority of practical applications.

From *figure 2b* you can see that the collector current exists even when the collector voltage is zero. This is due, in part, to the fact that when the potential barrier of the collector junction has zero potential applied to it, it attracts the minority carriers in the base region across the junction. To reduce the collector current to zero, a small direct positive collector voltage must be applied – as can be seen in *figure 2b*. The collector characteristics shown have been drawn in the third quadrant of a graph (zero point being in the top right here). This is to show that the collector junction is reverse biased. Normally output characteristics are drawn in the first quadrant of a graph (zero being the bottom left point).

Common emitter characteristics

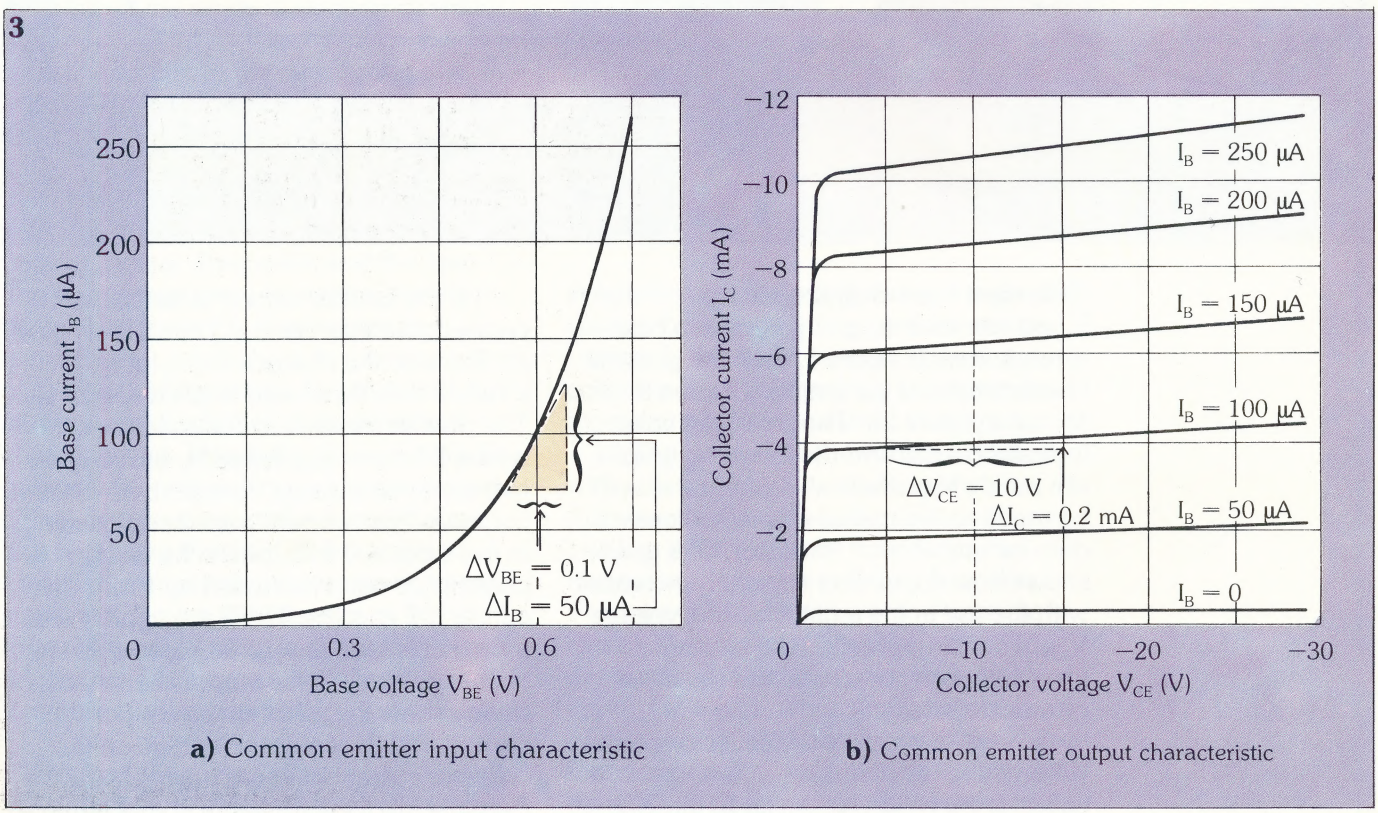
When a transistor is connected in common emitter mode, the input characteristic will show how the input current, which in this mode is the transistor's base current I_B, rises with the increase in the directly polarized base-emitter voltage V_{BE}. This characteristic is shown in *figure 3a* (it belongs to the same transistor used in *figure 2*).

The base current is the difference between the emitter and collector currents and will therefore be much smaller than the emitter current. The base-emitter voltage will, however, be the same as in the common base connection. If we consider Ohm's law:

$$R = \frac{V}{I}$$

and remember that the input current of the common emitter arrangement is lower than that of the common base mode, we can see that the input resistance of the common emitter must be higher than that of the common base.

(*continued in part 12*)



a) Common emitter input characteristic

b) Common emitter output characteristic

3. Common emitter characteristics: (a) input; **(b)** output.

TEST YOUR PROGRESS with the

I.T.E.C. QUIZ

FOLD 2

DIGITAL ELECTRONICS – 8

1. How does a real building block design usually differ from the sum-of-products design?

a There are usually fewer gates and fewer inputs to the gates
b Other kinds of gates may be substituted for AND gates and OR gates
c It's usually more economical
d All of the above

2. What is the largest number of data inputs a multiplexer can have, if it has two select inputs?

a Two
b Four
c Eight
d Any number

3. Which building block studied, best illustrates the principle of dividing the big processing tasks into a sequence of little ones?

a The decoder
b The multiplexer
c The demultiplexer
d The 4-bit adder

FOLD 3

4. Full-adders differ from half-adders because:

a Full-adders use less current
b Half-adders cannot accept a carry input
c Half-adders are faulty full-adders
d Full-adders operate at a higher speed

5. Positive and negative binary numbers are distinguished from each other by the use of a sign bit. When the sign bit is 1 the number is positive and when it is 0 the number is negative.

True or False?

6. If the binary numbers 1.101 and 0.101 are added together, the result is:

a 1.101
b 1.010
c 0.010
d 0.110

SOLID STATE – 10

1. The basic material used in the manufacture of most current semiconductor devices is:

a Monocrystalline silicon
b Grown-junction
c Gaseous boron
d n and p-type aluminium hydroxide
e b and d above

2. When monocrystalline silicon is grown, the seed is continuously lifted and both the seed and the crucible are continuously rotated to ensure uniform growth.

True or False?

3. Grown, alloyed and diffused junctions are the three general techniques for:

a Mechanical assembly of semiconductor products
b Electrical testing of semiconductor products
c Producing p-type and n-type regions in the semiconductor element of a device
d Application strictly to transistors
e a and b above

4. A big advantage of the diffusion process is that:

a Individual hand crafting of each separate semiconductor element is possible, thus achieving a high standard of quality
b It is a short one-step process
c Hundreds of semiconductor elements can be produced on the same slice in the time it would take to produce one
d It doesn't require intricate, precise techniques
e b and d above

Answers to last week's quiz

COMPUTER SCIENCE – 6	SOLID STATE – 9	DIGITAL ELECTRONICS – 7
1 f	1 b	1 b
2 c	2 True	2 c
3 True	3 e	3 b
		4 d
		5 c

FOLD 4

I enclose a cheque/postal order made payable to: GEJ Publishing Ltd. for a total of £ which I understand includes the cost of postage and packing.

N.B. Please allow 35 days for the delivery of both back numbers and binders.

When you have completed the order form, fill in your name and address in the space provided.

Then cut along the dotted line to detach the page, enclose your cheque/postal order, and fold the page carefully – following the instructions – to complete the reply paid envelope.
NO STAMP NECESSARY.

BLOCK CAPITALS PLEASE

If you already have an account number for I.T.E.C. binders please fill in here.

Complete the section below with one letter figure or space per square.

Mr / Mrs / Miss — Initials — Surname

Address

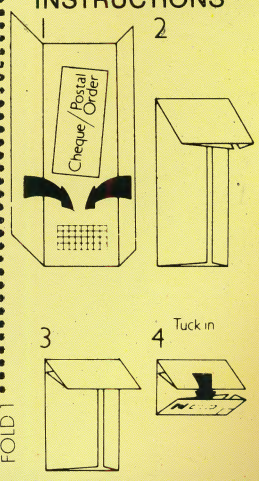
STD Code or Exchange name — Local Number

Telephone Number

FOLD 1

FOLD 4

FOLDING INSTRUCTIONS



How to order your binders...

These binders transform your copies into a handsome encyclopedia set – and give you maximum value from your home study course.

All you have to do to obtain your binders for *I.T.E.C.* is to tick the box(es) on the right and fill in your name and address on the section below. Enclose your cheque/postal order (made payable to G.E.J. Publishing Ltd). Fold carefully – following the instructions – to complete the reply paid envelope.
NO STAMP NECESSARY

If you tick both boxes on the right you will ensure that your future binders for *I.T.E.C.* arrive regularly, every ten issues. In this way you are saved the trouble of having to remember to order your binders individually and payment is simplified. There is no obligation on your part; you can cancel the standing order at any time.

How the binder set works: there are five binders in the complete set. You simply remove the covers from the weekly parts and store them, ten parts per binder plus the index part in the last binder.

☐ **Please send me my next binder now.**
I enclose cheque/postal order for £3.75 (including postage and packing).
If you require more than one binder to bring your collection up to date, please state how many binders you require and send cheque/postal order to cover the cost at £3.75 per binder.

☐ **I would also like to receive future binders as they are issued.**
I understand I will receive a payment advice for £3.75 (including postage and packing) with each binder, which I will pay only if I am absolutely satisfied with the binder. Otherwise I will return the binder and owe you nothing.

N.B. Please allow 35 days for delivery of your binders.

IMPORTANT:
Please read this carefully:

1. Do not complete this form if you have already asked for binders to be sent to you automatically as they are issued.

2. Readers in the Republic of Ireland: please remit your payment in sterling.

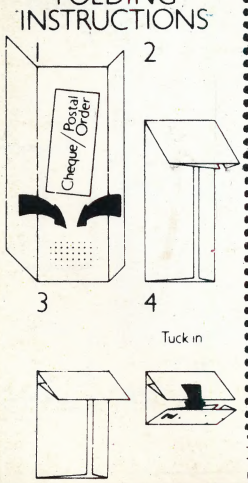
3. Readers not in UK or The Republic of Ireland, see inside cover for details of how to obtain binders.

Fold 2

Fold 3

FOLD 4

FOLDING INSTRUCTIONS



Fold 1

Postage will be paid by licensee

Do not affix Postage Stamps if posted in Gt. Britain, Channel Islands or N. Ireland.

BUSINESS REPLY SERVICE
Licence No. WD1106

2

G.E.J. Publishing Ltd (E.T.)
187 Oxford Street
London W1E 5EZ

Fold 1

FOLD 4

Guarantee

If you are not entirely satisfied with your binder, send it back immediately and it will be either exchanged, or, if you prefer, your money will be refunded in full.